

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 15:37:32 ; Search time 34.17 seconds
(without alignments)
2080.399 Million cell updates/sec

Title: US-09-905-056-292

Perfect score: 3362

Sequence: 1 MLNKMTHLPOQIMIGPRNR.....VHEPLLRMSKDNVQETQI 640

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A_Geneseq_032802.*
- 1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
 - 2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
 - 3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
 - 4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
 - 5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
 - 6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
 - 7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
 - 8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
 - 9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
 - 10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
 - 11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
 - 12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
 - 13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
 - 14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
 - 15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
 - 16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
 - 17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
 - 18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
 - 19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
 - 20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
 - 21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
 - 22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3362	100.0	640	20	AAV08100
2	3362	100.0	640	20	AAW85722
3	3362	100.0	640	20	AAV13394
4	3362	100.0	640	21	AAW24407
5	3362	100.0	640	21	AAV70673
6	3362	100.0	640	22	AAU12355
7	3362	100.0	640	22	AAU00826
8	3362	100.0	640	22	AAW80262
9	3362	100.0	640	22	AAW65292
10	3362	100.0	640	22	AAW53089
11	2131.5	63.4	713	22	AAE13006

best post-ant

12	1865	55.5	653	21	AAE24073	Human PRO1111 prot
13	1865	55.5	653	21	AAE66694	Membrane-bound pro
14	1865	55.5	653	22	AAE09438	Human sbgPRO331a p
15	1865	55.5	653	22	AAU12390	Human PRO1111 poly
16	1865	55.5	653	22	AAE65217	Human PRO1111 (UNQ
17	1856	55.2	653	20	AAE28806	cc339_4 secreted p
18	1838	54.7	606	22	AAU18035	Human immunoglobul
19	1832	54.5	694	21	AAE23033	Human SLIT protein
20	1710.5	50.9	590	21	AAE23034	Human SLIT protein
21	1634	48.6	553	22	ABE10349	Human cDNA SEQ ID
22	1561.5	46.4	448	21	AAE23044	Human SLIT protein
23	1164	34.6	441	21	AAE43091	Human ORFX ORF2855
24	974	29.0	302	21	AAE87066	Human secreted pro
25	974	29.0	302	21	AAE87066	Human gene 3 encod
26	845.5	25.1	281	22	AAE06043	Human cDNA SEQ ID
27	845.5	25.1	281	22	ABE10522	Human immunoglobul
28	688	20.5	224	21	AAU18084	Human secreted pro
29	688	20.5	224	21	AAE87135	Human gene 3 encod
30	586.5	17.4	614	20	AAE06112	Amino acid sequenc
31	583.5	17.4	620	22	AAW84596	Human membrane ass
32	575	17.1	606	22	AAE06799	Human neuronal gui
33	574.5	17.1	620	20	AAV13357	Amino acid sequenc
34	574.5	17.1	620	22	AAU12333	Human PRO227 poly
35	574.5	17.1	620	22	AAE80325	Human PRO227 prote
36	573.5	17.1	579	22	AAE06804	Mature human neuro
37	538.5	16.0	592	22	AAE09437	Human sbgTrango79a
38	487	14.5	548	22	AAE31161	Amino acid sequenc
39	481.5	14.3	766	22	AAE67505	Amino acid sequenc
40	466	13.9	1021	22	ABE60162	Drosophila melanog
41	446	13.3	771	22	ABE67512	Amino acid sequenc
42	438.5	13.0	713	20	AAV13385	Amino acid sequenc
43	438.5	13.0	713	22	AAE80253	Human PRO293 prote
44	435.5	13.0	785	21	AAE12448	Human hh00149 prot
45	435.5	13.0	789	21	AAE09968	Human brain-specif

ALIGNMENTS

RESULT 1
ID AAY08100 standard; Protein: 640 AA.
XX AC AAY08100;
XX AC AAY08100;
DT 11-SEP-2000 (first entry)
XX XX
DE Human PRO331 protein.

XX Inflammatory cell infiltration; immune response; T cell proliferation;
XX anti-inflammatory; anti-autoimmune; anti-diabetic; spondyloarthritis;
XX T cell-mediated disease; spondyloarthritis; sclerosis; renal disease;
XX inflammatory myopathy; hemolytic anemia; thrombocytopenia; thyroiditis;
XX diabetes mellitus; demyelinating polyneuropathy; Guillain-Barre syndrome;
XX multiple sclerosis; polyneuropathy; hepatitis; cirrhosis; enteropathy;
XX sclerosing cholangitis; inflammatory bowel disease; Whipple's disease;
XX skin disease; dermatitis; psoriasis; asthma; allergic rhinitis; tumor;
XX food hypersensitivity; urticaria; eosinophilic pneumonia; transplant;
XX idiopathic pulmonary fibrosis; graft rejection; PRO245; human;
XX PRO331.

OS Homo sapiens.
XX XX
XX WO9914241-A2.
XX XX
XX 25-MAR-1999.
XX XX
XX 17-SEP-1998; 98WO-US19437.
XX XX
XX 17-SEP-1997; 97US-0059119.
XX 18-SEP-1997; 97US-0059263.
XX 28-OCT-1997; 97US-0063550.
XX 12-NOV-1997; 97US-0065186.

PR 21-NOV-1997; 97US-0066364.
 PR 24-NOV-1997; 97US-0066770.
 PR 04-JUN-1998; 98US-0088026.
 XX (GETH) GENENTECH INC.
 XX
 PI Fong S, Goddard A, Gurney AL, Tumas D, Wood WI;
 XX
 DR WPI; 1999-229499/19.
 DR N-PSDB; AAX37738.
 XX
 XX Composition containing novel polypeptide PRO245, its agonist or
 PT antagonist -
 XX
 XX Example 1; Fig 33; 177pp; English.
 XX
 CC This invention describes a novel composition containing (apart from a
 CC carrier or excipient), a novel PRO245 polypeptide (I), its agonist or
 CC antagonist, or their fragments, for modulating: (i) infiltration of
 CC inflammatory cells into tissue; (ii) an immune response; or (iii) T cell
 CC proliferation. The composition increases or decreases any of the effects
 CC (i)-(iii). The products of the invention have anti-inflammatory,
 CC anti-autoimmune and anti-diabetic activity. (I), and its (ant)agonists
 CC and their fragments, are used to treat immune-related diseases,
 CC particularly T cell-mediated diseases. The diseases treated include
 CC systemic lupus erythematosus, rheumatoid arthritis, juvenile chronic
 CC arthritis, spondyloarthropathies, systemic sclerosis (scleroderma),
 CC idiopathic inflammatory myopathies (dermatomyositis, polymyositis),
 CC Sjogren's syndrome, systemic vasculitis, sarcoidosis, autoimmune
 CC hemolytic anemia (immune pancytopenia, paroxysmal nocturnal
 CC hemoglobinuria), autoimmune thrombocytopenia (idiopathic thrombocytopenic
 CC purpura immune-mediated thrombocytopenia), thyroiditis (Grave's disease,
 CC Hashimoto's thyroiditis, juvenile lymphocytic thyroiditis, atrophic
 CC thyroiditis), diabetes mellitus, immune-mediated renal disease
 CC (glomerulonephritis, tubulointerstitial nephritis), multiple sclerosis,
 CC idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome, chronic
 CC inflammatory demyelinating polyneuropathy, infectious hepatitis
 CC (hepatitis A, B, C, D, E and other non-hepatotropic viruses), autoimmune
 CC chronic active hepatitis, primary biliary cirrhosis, granulomatous
 CC hepatitis, and sclerosing cholangitis, inflammatory bowel disease
 CC (ulcerative colitis; Crohn's disease), gluten-sensitive enteropathy, and
 CC Whipple's disease. Autoimmune or immune-mediated skin diseases including
 CC bullous skin diseases, erythema multiforme, contact dermatitis, psoriasis,
 CC asthma, allergic rhinitis, atopic dermatitis, food hypersensitivity,
 CC urticaria, eosinophilic pneumonia, idiopathic pulmonary fibrosis,
 CC hypersensitivity pneumonitis, and transplantation associated diseases
 CC (graft rejection, and graft-versus-host-disease). (I), its (ant)agonists
 CC or fragment can also be used as an adjuvant in treatment of tumors.
 CC Antibodies against (I) can also be used for diagnosing such diseases.
 CC This sequence represents the human PRO331 protein which is described in
 CC the method of the invention.
 XX
 XX Sequence 640 AA;

Query Match 100.0%; Score 3362; DB 20; Length 640;
 Best Local Similarity 100.0%; Pred. No. 2.3e-234;
 Matches 640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLNKMTHLPQOIMIGPRFNALFDPPLLVLLALQLLVVAGLVRAQTCPSCVCSNQFSKV 60
 DB 1 mLnkmtlhpdqimigprfnalfdpllvllalqlvvaglvraqtcpsvcscsnqfskv 60
 QY 61 ICVRKRLREVPDGTSTRLNLHENOIQIKVNSFKHLRHLLETLQLSRNHRTIEIGAF 120
 DB 61 icvrknlrevpdgdsttrlnlnhenoqiikvnsfkhlrlhlleqlslsrnhrtieigaf 120
 QY 121 NGLANLTLELDFNRLTTPNGAFVYLSKLKELWLRNPIESTPSYAFNRIPSLRRLDLG 180
 DB 121 nglanltlelfdnrlttpngafvylskelwlrnpiestpsyafnrslrldlg 180
 QY 181 ELKRLSYISEGAFGLSNRLRYLNAMCNLREIPNLTLPLIKLDELDSGNLSAIRPGSFQ 240
 DB 181 elkrslsyisegafglslnrlrylnamcnlreipnltplikldeidsgnlsairpgsfq 240

DB 181 elkrslsyisegafglslnrlrylnamcnlreipnltplikldeidsgnlsairpgsfq 240
 QY 241 GLMHLQKLWMTQSOIQVTERNAFDNLOSLEINLAHNNLTLLPHDLFTPLHLERHLHH 300
 DB 241 glmhlqklwmqsqiqvternafdnlslelnlahnnltllphdlftplhlhlerhlhh 300
 QY 301 NPWNCNDIILWLSWIKDMAFSPNTACCARCNTPPNLKGRYIGELDONFTCYAPVIVEPP 360
 DB 301 npwncncdiilwlswwikdmapsnaccarcntppnlkgrylgeidqnyftcyapvivepp 360
 QY 361 ADLNVTEGMAELKCRASSTLSVTSWITPNTGVTWTHGAYKVRIAVLSOGTLNFTNVTVD 420
 DB 361 adlnvtegmaaelkcrasstlsvtswtipngtvmthgaykvrilavlsdgtlnftnvtvd 420
 QY 421 TGMVTCVNSVGNVTASATLNVTAATTPSFYSTVETMETPSODEARTDDNNGVPTP 480
 DB 421 tgmvtcvnsvngnvtasatlnvtaattpsfstvctvemetpsqdeartddnngvptp 480
 QY 481 VVDWETTNVTSLTPPQSTREKFTTIPVTDINSIGPIDEVMKTKIIIGCFVAITLMA 540
 DB 481 vvdwettntvtsltpqstrestekfttvpvtdinsigpidevmkttkiiigcfvaitlma 540
 QY 541 AVMLVIFYKKRKHRRQNHRRHAPTRFVELIINVDDITGDTPMESHLPMPAIEHEHLNHYS 600
 DB 541 avmlvifykmrkqhrqhnhaptrtvelinvdditgdtpmeshlmpalehehlnhyns 600
 QY 601 YKSPFNHTTTVTNTINSIHSSVHEPILLIRMSKDNVQETQI 640
 DB 601 ykspfntttvtntinsihssvhepillsirmskdnvqetqi 640

RESULT 2

AAW85722
 ID AAW85722 standard; Protein; 640 AA.
 XX
 AC AAW85722;
 XX
 DT 27-SEP-1999 (first entry)
 XX
 DE Novel protein (Clone AS209_1).
 XX
 KW Polynucleotide; protein; nutrition; cytokine; cell proliferation;
 KW cell differentiation; immunostimulation; immunosuppression;
 KW haematopoiesis regulation; tissue growth; activin; inhibin;
 KW chemotaxis; chemokinesis; haemostasis; thrombolysis; receptor;
 KW ligand; anti-inflammatory; tumour suppression; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO9920644-A1.
 XX
 PD 29-APR-1999.
 XX
 PF 16-OCT-1998; 98WO-US22034.
 XX
 PR 18-OCT-1997; 97US-0955557.
 XX
 PA (GEMY) GENETICS INST INC.
 XX
 PI Agostino MJ, Bowman MR, Evans C, Jacobs K, Lavallie ER;
 PI McCoy JM, Merberg D, Racie LA, Spaulding V, Treacy M;
 DR WPI; 1999-288272/24.
 DR N-PSDB; AAX08687.
 XX
 PT New polynucleotides encoding secreted human proteins
 XX
 PS Claim 26; Page 109-111; 136pp; English.
 XX
 CC The new human secreted proteins are encoded by polynucleotides
 CC obtained from human placenta, adult testes, fetal kidney, fetal
 CC brain, adult brain, adult blood and adult cDNA libraries.
 CC The polynucleotides and proteins are predicted to have biological

CC activities which would make them suitable for treating, preventing or
CC ameliorating medical conditions in humans and animals. Suggested
CC activities include nutritional activity, cytokine and cell
CC proliferation/differentiation activity, immune stimulating (e.g. as
CC vaccines) or suppressing activity, haematopoiesis regulating
CC activity, tissue growth activity, activin/inhibin activity,
CC chemotactic/chemokinetic activity, haemostatic and thrombolytic
CC activity, receptor/ligand activity, anti-inflammatory activity,
CC cadherin/tumour invasion suppressor activity, and tumour inhibition
CC activity. The polynucleotides are also stated to be useful for gene
CC therapy. The sequences identified by a secretory leader
CC sequence motif in the polynucleotide and it is thought that the
CC encoded proteins have biological activity by virtue of their secreted
CC nature. This polypeptide was encoded by a clone designated AS209_1
CC (See AAX08687).

XX Sequence 640 AA;

Query Match 100.0%; Score 3362; DB 20; Length 640;
Best Local Similarity 100.0%; Pred. No. 2.3e-234;
Matches 640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLNMTLHPQIMIGPRNRLFPDLLVLLALQLLVVAGLVRAQTCPSCVCSNQPSKV 60
Db 1 mlnkmtlhpqimigprnrlfddlvlvllalqlvvaglvraqtcpsvcscsqfkskv 60

QY 61 ICVRNLRVDPDGTSTNRLNLHENQIQIKVNSFKHLRHLLEILQLSRNHRITIEIGAF 120
Db 61 icvrnlrevpdgstnrlnlhnenqiiqikvnsfkhlrhlleilqlsrnhritleigaf 120

QY 121 NGLANLTLFDRNLRTTIPNGAFVYLSKLKELWLRNPIESYAFNRIPSLRRLDLG 180
Db 121 nglanlntleifdrnlrtticipngafvylskelwlrnpiesiyafnrslrrldlg 180

QY 181 ELKRLSYISEGAFGLSLNRYLNLAACNLRETPNLPLIKLDELDSLNSHLSAIRPGSFQ 240
Db 181 elkrlysiisegafglslnrylnlamcnlreipnlpllikdeldslnsghlsairpgsfq 240

QY 241 GLMHQLKLMWISQIQVIERNAFNLQSLVEINLAHNNLTLLPHDLFTPLHLHLRIHLRH 300
Db 241 glmhqlklwmisqiqviernafndnlqslveinlahnnltllphdlftplhlhlrihlrh 300

QY 301 NPWNCNCDILNLSWIKDMASNTACCARNTPPNLKGRIYIGELDQNYFTCYAPVIVEPP 360
Db 301 npwncncdilnlswwikdmasntaccarnctppnlkgriyigeidqnyftcyapvivepp 360

QY 361 ADLNVTEGMAELKCRASLTSTSVSWITPNCVTMTHGAYKVRVIAVLSDGTINFTVTQD 420
Db 361 adlnvtegmaaelkcrasltstsvswitpncvtmthgaykvrviavlsdgtinfntvtqd 420

QY 421 TGMVTCMYSNVGNVTASATLNVAATTPFSYFTSVTVMERSQDEARTTDNNVGTP 480
Db 421 tgmvtcmysnvgntasatlnvtaattppfsyftsvtvmersqdearttdnnvgtp 480

QY 481 VVDNETNVTSLPQSTRTEKFTTIPVTDINSIPGIDVEMKTKIIGCFVAITLMA 540
Db 481 vvdnetnvtsltpqstrtektfttvpvtdinsipgidvemtktliigcfvaitlma 540.

QY 541 AVMLVIFYKMRKHQRHQRHAPTRVETLINVDDDETGTDPMESHLPMPAIEHNLHNS 600
Db 541 avmlvifykmrkhqrhqrhaptrevetlinvdddetgtatpmeshlmpaiehehlnhns 600

QY 601 YKSPFNHTTNTINSIHSSVHEPLLRMSKDNVQETQI 640
Db 601 ykspfnhtttntinsihssvhepllrmskdnvqetqi 640

RESULT 3
AAY13394
ID AAY13394 standard; Protein; 640 AA.
XX
AC AAY13394;

XX 25-JUN-1999 (first entry)
DT
XX Amino acid sequence of protein PRO331.
DE
XX Secreted protein; transmembrane protein; human; enterocolitis;
KW Zollinger-Ellison syndrome; gastrointestinal ulceration;
KW congenital microvillus atrophy; skin disease; cell growth;
KW abnormal keratinocyte differentiation; psoriasis; epithelial cancer;
KW Parkinson's disease; Alzheimer's disease; ALS; neuropathy;
KW fibromodulin; dermal scarring; Usher Syndrome; Atrophia areata;
KW anti-thrombotic; wound healing; tissue repair.
XX
OS Homo sapiens.
XX
XX WO9914328-A2.
PN
XX 25-MAR-1999.
PD
XX 16-SEP-1998; 98WO-US19330.
PF
XX 25-NOV-1997; 97US-0066840.
PR 17-SEP-1997; 97US-0059113.
PR 17-SEP-1997; 97US-0059115.
PR 17-SEP-1997; 97US-0059117.
PR 17-SEP-1997; 97US-0059119.
PR 17-SEP-1997; 97US-0059121.
PR 17-SEP-1997; 97US-0059122.
PR 17-SEP-1997; 97US-0059184.
PR 18-SEP-1997; 97US-0059263.
PR 18-SEP-1997; 97US-0059266.
PR 15-OCT-1997; 97US-0062125.
PR 17-OCT-1997; 97US-0062285.
PR 17-OCT-1997; 97US-0062287.
PR 21-OCT-1997; 97US-0063496.
PR 24-OCT-1997; 97US-0062814.
PR 24-OCT-1997; 97US-0062816.
PR 24-OCT-1997; 97US-0063045.
PR 24-OCT-1997; 97US-0063120.
PR 24-OCT-1997; 97US-0063121.
PR 24-OCT-1997; 97US-0063127.
PR 27-OCT-1997; 97US-0063128.
PR 27-OCT-1997; 97US-0063329.
PR 27-OCT-1997; 97US-0063327.
PR 28-OCT-1997; 97US-0063541.
PR 28-OCT-1997; 97US-0063542.
PR 28-OCT-1997; 97US-0063544.
PR 28-OCT-1997; 97US-0063549.
PR 28-OCT-1997; 97US-0063550.
PR 28-OCT-1997; 97US-0063564.
PR 29-OCT-1997; 97US-0063435.
PR 29-OCT-1997; 97US-0063704.
PR 29-OCT-1997; 97US-0063732.
PR 29-OCT-1997; 97US-0063738.
PR 29-OCT-1997; 97US-0063734.
PR 29-OCT-1997; 97US-0064215.
PR 31-OCT-1997; 97US-0063870.
PR 31-OCT-1997; 97US-0064103.
PR 03-NOV-1997; 97US-0064248.
PR 07-NOV-1997; 97US-0064809.
PR 12-NOV-1997; 97US-0065186.
PR 17-NOV-1997; 97US-0065846.
PR 18-NOV-1997; 97US-0065693.
PR 21-NOV-1997; 97US-0066120.
PR 21-NOV-1997; 97US-0066364.
PR 24-NOV-1997; 97US-0066772.
PR 24-NOV-1997; 97US-0066466.
PR 24-NOV-1997; 97US-0066770.
PR 24-NOV-1997; 97US-0066511.
PR 24-NOV-1997; 97US-0066453.
XX
XX (GETH) GENENTECH INC.

Chen J, Goddard A, Gurney AL, Pennica D, Wood WI, Yuan J;
WPI: 1999-229533/19.
N-ESDB: AAX52265.
New isolated human genes and polypeptides used in, e.g. treatment of
gastrointestinal ulceration
Claim 12; Fig 104; 320pp; English.
AAV13344-403 represent secreted and transmembrane human proteins.
The cDNA sequences are obtained from cDNA libraries, prepared from
fetal lung, fetal kidney, fetal brain, fetal liver and fetal retina.
The encoded polypeptides have specific uses based on their homology to
known polypeptides, e.g. PRO211 and PRO217 can be used for disorders
associated with the preservation and maintenance of gastrointestinal
mucosa and the repair of acute and chronic mucosal lesions
(e.g. enterocolitis, Zollinger-Ellison syndrome, gastrointestinal
ulceration and congenital microvillus atrophy), skin diseases associated
with abnormal keratinocyte differentiation (e.g. psoriasis, epithelial
cancers such as lung squamous cell carcinoma of the vulva and gliomas),
potent effects on cell growth and development, diseases related to
growth or survival of nerve cells including Parkinson's disease,
Alzheimer's disease, ALS, neuropathies or cancer. PRO265 can be used as
a target for anti-tumor drugs. PRO533 may be used in the treatment
of Usher Syndrome or Atrophia areata; PRO269 can be used as an
anti-thrombotic agent; PRO287 polypeptides and portions may have
therapeutic applications in wound healing and tissue repair; PRO317 can
be used for treating problems of the kidney, uterus, endometrium, blood
vessels, or related tissue, e.g. in the heart of genital tract.

Query Match 100.0%; Score 3362; DB 20; Length 640;
Best Local Similarity 100.0%; Pred. No. 2.3e-234;
Matches 640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLNKLTPQOITGIPRNFALFDPLLVLLALQLLVAGLVRAQTCPSCVCSNQFSKV 60
DB 1 mLnKmltpqqimqprnfalfdpllvllalqlvvaglvraqtcpsvcscnqfskv 60
QY 61 ICVRKNLREVPDGSITNRLNLHENQIQIKVNSFKHLRLHLEIQLSRNIRTEIGAF 120
DB 61 icvrknlrevpdgsitnrlnlhenqiiqvnsfkhlrlhleilqlsrnirtieigaf 120
QY 121 NGLANLTLFLDNLRTTPNGAFVYLSKLKELWLRNPNIESIPSYAFNRIPSLRLDLG 180
DB 121 nglanltlflndrlttpngafvylsklkelwlrnnpiesipsyafnrpslrlldlg 180
QY 181 ELKRLSYISEGAFGLSNLRYLNAMCNLRTPNLTPLIKLDELDSLGNHLSAIRPGSFQ 240
DB 181 elkrlyisegafglsnlrylnamcnlrtpnltplikdeldslgnhlsairpgsfq 240
QY 241 GLMLHQLKWLMTQSQTQVIERNAFQNLQSLVEINLAHNNLTLLPHDLFTPLHHLRIHLRH 300
DB 241 glmlhqlkwlmsqslqviernafndqlslveinlahnnltllphdlftplhhlrihlrh 300
QY 301 NPWNCNCDILMSWIKDMASNTACCARNTPPNLKGRYGELDQNYFTCYAPVIVEPP 360
DB 301 npwncncdilwswikdmassntaccarnctppnlkgrygeldqnyftcyapvivepp 360
QY 361 ADLNVTEGMAELKCRASLTSTVSWITPNCVTMGHAYKVRVIAVLSGTLNFTNTVQD 420
DB 361 adlnvtegmaaelkcrasltstsvswitpncvtmgghaykvrviavlsdgtlnftntvqd 420
QY 421 TGMVTCWVSNVGNFTASATLNVAATTFPFSYFSTVVTMEPSQDEARTDNNVGTP 480
DB 421 tgmvtcmvsnvgnftasatlnvtaattfpfsyfstvtvetmepsqdeartdnnvgtp 480
QY 481 VDWETTNNVTSLPQSTRSTEKFTTIPVTDINSIGIPIDEVMKTKIIIGCFVAITLMA 540

Db 481 vdwettntvtsltpqstrstektftipvtinsigipidevmktkiiigcfvaitlma 540
QY 541 AVMLVIFYKMRKHRRONHAPTRTVELINVDDETGTGTPMESHLPMPAIEHEHLNHYNS 600
DB 541 avmlvifykmrkhrqnqhapttrtvelinvddeitgtpmeshlpmpalenehlnhyns 600
QY 601 YKSPFNHTTNTVNTINSIHSSVHEPLLIRMSKDNVQETQI 640
DB 601 ykspfnttntvntinsihssvhpepllirmnskdnvqetqi 640
RESULT 4
AAB24407
ID AAB24407 standard; Protein; 640 AA.
XX AAB24407;
XX AAB24407;
DT 07-NOV-2000 (first entry)
DE Human PRO331 protein sequence SEQ ID NO:107.
KW Human: PRO; promotion; inhibition; angiogenesis; cardiovascularisation;
KW diagnosis; trauma; wound; cancer; atherosclerosis; cardiac hypertrophy;
KW angiogenic; proliferative; cardiac; cardiovascular; antiatherosclerotic;
KW cytostatic; gene therapy; vaccine.
XX Homo sapiens.
XX WO200032221-A2.
XX 08-JUN-2000.
XX 30-NOV-1999; 99WO-US28313.
XX 01-DEC-1998; 98WO-US25108.
XX 16-DEC-1998; 98US-0112850.
XX 12-JAN-1999; 99US-0115554.
XX 08-MAR-1999; 99WO-US05028.
XX 12-MAR-1999; 99US-0123957.
XX 28-APR-1999; 99US-0131445.
XX 14-MAY-1999; 99US-0134287.
XX 02-JUN-1999; 99WO-US12252.
XX 20-JUL-1999; 99US-0144758.
XX 26-JUL-1999; 99US-0145698.
XX 01-SEP-1999; 99WO-US20111.
XX 08-SEP-1999; 99WO-US20594.
XX 13-SEP-1999; 99WO-US20944.
XX 15-SEP-1999; 99WO-US21090.
XX 05-OCT-1999; 99WO-US21547.
XX 29-OCT-1999; 99WO-US23089.
XX 99US-0162506.
XX (GETH) GENENTECH INC.
XX Ashkenazi AJ, Baker KP, Ferrara N, Gerber H, Hillan KJ, Goddard A;
PI Godowski PJ, Gurney AL, Klein RD, Kuo SS, Paoni NF, Smith V;
PI Watanabe CK, Williams PM, Wood WI;
XX WPI; 2000-412154/35.
XX N-ESDB; AAA77596.
XX Nucleic acids encoding PRO polypeptides useful for preventing,
PT diagnosing and treating disorders a cardiovascular, endothelial or
PT angiogenic disorders in mammals -
XX Claim 72; Fig 40; 315pp; English.
XX The present invention describes nucleic acids encoding PRO polypeptides
CC useful for preventing, diagnosing and treating disorders a
CC cardiovascular, endothelial or angiogenic disorder in mammals by
CC modulating cell proliferation, angiogenesis and cardiovascularisation,

CC and for identifying agonists and antagonists of these processes. The
CC nucleic acids and the proteins they encode may be used in the
CC prevention, treatment and diagnosis of diseases associated with
CC inappropriate PRO expression such as cardiovascular, endothelial or
CC angiogenic disorders in mammals (e.g. atherosclerosis, cancers and
CC cardiac hypertrophy). For example, the nucleic acids (NCs) and vectors
CC containing them and the PRO polypeptide may be used to treat disorders
CC associated with decreased PRO expression. AAA77510 to AAA77721 and
CC AAB24388 to AAB24435 represent nucleotide and protein sequences used in
CC the exemplification of the present invention.

XX SQ Sequence 640 AA;

```
Query Match          100.0%; Score 3362; DB 21; Length 640;
Best Local Similarity 100.0%; Pred. No. 2.3e-234;
Matches 640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLNKMTHLPQIMIGPRFNRAFDPLLVLLALQLLVVAGLVRAQTCPSVCSNQSFKV 60
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 mlnkmtlhpqqimigprfnralfdpallvllalqlvvaglvraqtcpsvcscnqfskv 60

Qy 61 ICVRKNREVDPGISTNTRLNLHENQIIKVSFKHLRHLLEIQLSRNHIRIEICAF 120
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 icvrknrevdpdgistntrlnlhengiqikvnsfkhlrhllelqlsrnhirtieigaf 120

Qy 121 NGLANLTLELFNRLATIPNGAFVYLSKKELWLRNNPIESIPSNRIPSLRRLLDLG 180
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 nglanltlelfndrltippngafvylskelwlrnnpiesipsyafnrpslrrldlg 180

Qy 181 ELKRLSYISEGAFGLSNRLYLNAMCNLRIPNLTPLIKDELDLSGNHLSAIRPGSFQ 240
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 181 elkrlyisegafeglsnrllynlamcnlreipnltplikideidlsghnhsairpgsfq 240

Qy 241 GLMHLOKLWMIQSOTQVTERNAFNQSLVEINLAHNNLTLLPHDLFTPLHLERHLHH 300
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 241 glmhloklwmiqsoqvternafnqslveinlahnnltllphdlftplhlhlerhlh 300

Qy 301 NPWNCNDILSWWIKDMAPSNTACCARNTPPNLKGRIYIGELDONFTCYAPVIVEPP 360
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 301 npwncndilswwikdmapsntaccarnppnlkryigeldqnyftcyapvivepp 360

Qy 361 ADLNVTEGMAELKCRASLTLSVSWITPNGTVTHGAYKRVIAVLSGDTLNFNTVQD 420
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 361 adlnvtegmaaelkcrasltsvswitpngtvthgaykryviavlsdgtlnftnvtq 420

Qy 421 TGMVTCMYSNMGNTASATLNVAATTPFSYFTSTVETMEPSQDEARTDDNNVGPTP 480
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 421 tgmvtcmysnmgntasatlnvtaattptfsyfstvtvetmepsqdeartddmngptp 480

Qy 481 VVDWETTNVTTSLTPQSTRSTEKFTTIPVTDINSIGIPGIDEVMKTKIIGCFVAITLMA 540
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 481 vvdwettntvttsltpqstrstekfttvpvtdinsigipgidevmkttkiigcfvaitlma 540

Qy 541 AVMLVIFYKMRKHQRHQRHAPTRVTEIINVDDDEITGTPWESHLPMPAIEHEHLNHYNS 600
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 541 avmlvifykmrkqhrrqhrrhaptvteiinvddeitgtpmeshlpmpaiehehlnhyns 600

Qy 601 YKSPENHTTNTVNTINSIHSSVHEPLLIRMSKDNVQETQI 640
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 601 ykspfntttntvntinsihssvhepllirmnskdnvqetqi 640

RESULT 5
AAY70673
ID AAY70673 standard; Protein; 640 AA.
XX
AC AAY70673;
XX
DT 18-JUL-2000 (first entry)
XX
DE Human PRO331 protein.
XX
```

KW PRO331; UNQ292; dermatological; immunosuppressive; antiinflammatory;
KW immunostimulant; antiasthmatic; antirheumatic; antiarthritic; virucide;
KW antiallergic; haemostatic; hepatotropic; antidiabetic; antianaemic;
KW nephrotropic; neuroprotective; anticoagulant; immunological disorder;
KW lung; pneumonia; skin; psoriasis; kidney; glomerulonephritis; arthritis;
KW spondyloarthropathy; SLE; systemic lupus erythematosus; scleroderma;
KW idiopathic inflammatory myopathy; anaemia; thrombocytopenia; diabetes;
KW thyroiditis; Grave's disease; demyelinating disease; multiple sclerosis;
KW Crohn's disease; hepatobiliary disease; hepatitis; asthma; human;
KW graft-versus-host-disease.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 40..46 /note= "N-myristoylation site"
FT Modified-site 73..79 /note= "N-myristoylation site"
FT Modified-site 118..124 /note= "N-myristoylation site"
FT Modified-site 183..187 /note= "cAMP and cGMP-dependent protein kinase phosphorylation site"
FT Modified-site 191..197 /note= "N-myristoylation site"
FT Modified-site 228..234 /note= "N-myristoylation site"
FT Modified-site 237..243 /note= "N-myristoylation site"
FT Modified-site 268..272 /note= "N-myristoylation site"
FT Modified-site 278..282 /note= "Casein Kinase II phosphorylation site"
FT Modified-site 364..368 /note= "N-glycosylation site"
FT Modified-site 390..394 /note= "N-glycosylation site"
FT Modified-site 391..397 /note= "N-glycosylation site"
FT Modified-site 412..416 /note= "N-myristoylation site"
FT Modified-site 415..419 /note= "N-glycosylation site"
FT Modified-site 417..421 /note= "N-glycosylation site"
FT Modified-site 422..428 /note= "Casein Kinase II phosphorylation site"
FT Modified-site 433..439 /note= "N-myristoylation site"
FT Modified-site 434..438 /note= "N-myristoylation site"
FT Modified-site 442..446 /note= "N-glycosylation site"
FT Modified-site 465..469 /note= "N-glycosylation site"
FT Modified-site 488..492 /note= "Casein Kinase II phosphorylation site"
FT Modified-site 531..537 /note= "N-glycosylation site"
FT Modified-site 579..583 /note= "N-myristoylation site"
FT Modified-site 606..610 /note= "Casein Kinase II phosphorylation site"
FT Modified-site 620..624 /note= "N-glycosylation site"
FT Modified-site /note= "Casein Kinase II phosphorylation site"
XX WO200015797-A2.
XX
XX 23-MAR-2000.
XX 15-SEP-1999; 99WO-US21547.
XX 17-SEP-1998; 98US-0100858.
PR

```
PR 17-SEP-1998; 98WO-US19437.
XX (GETH ) GENENTECH INC.
XX Fong S, Goddard A, Gurney AL, Tumas D, Wood WT;
XX WPI; 2000-271435/23.
DR N-PSDB; AAZ52207.
XX
PT Composition for treatment and diagnosis of immune related diseases e.g.
PT Grave's disease comprises a PRO245, PRO217, PRO301, PRO266, PRO335,
PT PRO331 or PRO326 polypeptide or its agonists or antagonists (preferably
PT antibodies) -
XX
XX Example 1; Fig 14; 201pp; English.
XX
CC The present sequence is the human protein PRO331, encoded by UNQ292 cDNA,
CC designated as clone DNA40981. It is isolated from human foetal brain
CC tissue. Portions of PRO331 has homology to the LIG-1 protein.
CC It enhances or suppresses the infiltration of inflammatory cells into
CC tissues, proliferation of T-lymphocytes and modulates the immune
CC response. This sequence is useful for treatment of immune related
CC disorders, like SLE, rheumatoid/juvenile arthritis, spondyloarthritis,
CC systemic sclerosis (scleroderma), idiopathic inflammatory myopathies such
CC as dermatomyositis, Sjogren's syndrome, systemic vasculitis, sarcoidosis,
CC autoimmune haemolytic anaemia, thrombocytopenia, thyroiditis e.g. Grave's
CC disease, diabetes mellitus, immune-mediated renal disease e.g.
CC glomerulonephritis, demyelinating diseases such as multiple sclerosis and
CC Guillain-Barre syndrome, hepatobiliary diseases like hepatitis and
CC primary biliary cirrhosis, inflammatory and fibrotic lung diseases such
CC as inflammatory bowel disease (e.g. Crohn's disease), autoimmune or
CC immune-mediated skin diseases such as psoriasis, allergies like asthma,
CC immunological diseases of the lungs such as eosinophilic pneumonia and
CC transplantation associated diseases such as graft-versus-host-disease.
XX
XX Sequence 640 AA;
```

```
Query Match 100.0%; Score 3362; DB 21; Length 640;
Best Local Similarity 100.0%; Pred. No. 2,3e-234;
Matches 640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 MLNKLTHPQIMIGPRNRFDPDLVLLALQLLVVAGLVRAQTCPVCSQNSQSKV 60
Db 1 mlNkmtlhpgqimigprfnralfdpllvllalqlvvaglvraqtcpsvcscnqfskv 60
Qy 61 ICVRKNLREVDPDGTSTNRLNLHENOQIIKVNSEKHLRLLEILQLSRNHRITIEIGAF 120
Db 61 icvrknlrevdpdgstntrlnlhngeniqlikvnsfkhlrhlleqlsrnhritieigaf 120
Qy 121 NGLANLTLFELFDRNLRTIPNGAFVYLSKELWLRNNPIESIPYAFNRIPSLRRDLG 180
Db 121 nglanlntlelfdnrtrtippngafvylskelwlrnnpiesipynrripsrlridlg 180
Qy 181 ELKRLSYISEGAFGLSLNRLVNLAMCNLRIPNLPLIKLDELDSLGNHLSAIRPGSFQ 240
Db 181 elkrlyisegafglslnrlvnlamcnlreipnlpllikdeldslgnhlsairpgsfq 240
Qy 241 GLMHLQKLWMTQSQIQVIERNAFNLQSLVEINLAHNLTLLPHDLFTPLHLHLRIHLHH 300
Db 241 glmhqlklwmqlsqiqviernaflnqlslveinlahnltllphdlftplhlhlrihlhh 300
Qy 301 NPWNCNDILWLSWIKDMAPSNTACCARCNTPPNLKGRYTGELDQNYFTCYAPVIVEPP 360
Db 301 npwncndilwlswwikdmapsntaccarcntppnlkgrylgelydnyftcyapvivepp 360
Qy 361 ADLNVTEGMAELKCRASSTLSVSWITPNCVTMGTHGAKVRIAVLSDTLNFNTVIVQD 420
Db 361 adlnvtegmaaelkcrastsitsvswitpncvtmgthgavkriavlsdgtlnfncvntvqd 420
Qy 421 TGMVTCMVSNGVNTASATLNVAATTPPSYFSTVVTETMPSQDEARTDNNVGTP 480
Db 421 tgmvtcmvsvngvntasatlnvtaattppsfystvtvetmepsqdeartdnnvgtp 480
```

RESULT 6

```
AAU12355
ID AAU12355 standard; Protein; 640 AA.
XX
AC AAU12355;
XX
DT 24-OCT-2001 (first entry)
XX
DE Human PRO331 polypeptide sequence.
XX
KW Human secretory and transmembrane; PRO; mammalian; cancer; lung;
KW breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;
KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
KW adipocyte; A-peptide; factor VIIA; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200140466-A2.
XX
PD 07-JUN-2001.
XX
PF 01-DEC-2000; 2000WO-US32678.
XX
PR 01-DEC-1999; 99WO-US28301.
PR 01-DEC-1999; 99WO-US28634.
PR 02-DEC-1999; 99WO-US28551.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.
PR 09-DEC-1999; 99US-0170262.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 30-DEC-1999; 99WO-US31243.
PR 06-JAN-2000; 2000WO-US00277.
PR 06-JAN-2000; 2000WO-US00376.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US04914.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 20-MAR-2000; 2000WO-US07377.
PR 21-MAR-2000; 2000WO-US07532.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 10-NOV-2000; 2000WO-US30873.
XX
PA (GETH ) GENENTECH INC.
XX
PI Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX
XX WPI; 2001-408281/43.
XX
XX N-PSDB; AAS21427.
```

xx Isolated, secretory and transmembrane PRO polypeptide used to detect
PT other PRO polypeptides, link bioactive molecules to cells expressing
PT PRO polypeptides, and detect the presence of mammalian tumours e.g.
PT lung, breast, prostate, cervical
xx
PS Claim 12: Fig 368; 813pp; English.
xx
CC AAU12172-AAU12446 represent novel human secretory and transmembrane
CC PRO polypeptides. The PRO polypeptides are useful to detect other
CC PRO polypeptides, to link bioactive molecules to cells expressing
CC PRO polypeptides, to modulate biological activities of cells expressing
CC PRO polypeptides, and to detect the presence of mammalian lung, colon,
CC breast, prostate, rectal, cervical or liver tumours by comparing PRO
CC polypeptide expression in a cell sample to that in a control sample.
CC Some of the 275 sequences are also useful to stimulate the release of
CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the
CC proliferation or differentiation of chondrocytes, the proliferation or
CC gene expression in pericyte cells, the release of proteoglycans from
CC cartilage, the proliferation of inner ear utricular supporting cells or
CC of T-lymphocytes, the release of a cytokine from peripheral blood
CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of
CC the PRO polypeptides may modulate glucose or free fatty acid uptake by
CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
CC to factor VIIA. The PRO polypeptides can be used in assays to identify
CC molecules involved in binding interactions. The polynucleotides encoding
CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,
CC transgenic or knock out animals and can be used in gene therapy.
xx
SQ Sequence 640 AA;

Query Match 100.0%; Score 3362; DB 22; Length 640;
Best Local Similarity 100.0%; Pred. No. 2.3e-234;
Matches 640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLNKLTHPQIMIGPRNRLFPDLLVLLALQLLVVAGLVRAQTCPSCVCSNQSKV 60
DB 1 mlnkmtlhpqimigprnrafdpllvllalqlvvaglvraqtcpsvcscsnqskv 60
QY 61 ICVRKNREVPGDSTNFRNLNHNQIIVKNSFKHLRHLLETLQLSNRHRIEIGAF 120
DB 61 icvrknrevpgdstnfrnlhnenqiiqkvnshkrlhrllelqlsrnhrietiagf 120
QY 121 NGLANLTLLEFDNRLFTIPNGAFVYISKLKELWLRNPIESIPSNRIPSLRRDLG 180
DB 121 nglanltlefdnrlftltpngafvysklkelwlrnpiessyafnrripsrrldlg 180
QY 181 ELKRLSYISEGAFEGLSNLRVNLNLMWNLREIPNLPLIKDELDSGNHLSAIRPGSFQ 240
DB 181 elkrlyisegafeglsnlynlamcnlreipnltplikidelsgnhsairpgsfq 240
QY 241 GLMHLOKLWMTQSOIQVTERNAFNLQSLVEINLAHNLTLLPHDLFTPLHLHRIHLH 300
DB 241 glmhqlklwmtqslqvternafnqlsglveinlahnltlphdlftplhlhlerihlh 300
QY 301 NPWNCNDILWSWIKMDAPSNNTACCARNTPPNLKGRIYIGELDONFTCYAPVIVEPP 360
DB 301 npwncndilwswikmdapsntaccarntpbnlkgryigeldqnyftcyavivepp 360
QY 361 ADLNVTEGMAELKCRASSTLSVSWITPNTGVTWTHGAYKRVIAVLSGDTLNFNTVQD 420
DB 361 adlnvtegmaaelkcrasstlsvswitpntgvtwmthgaykryriavlsdgtlnfntvq 420
QY 421 TGMVTCVNSVNGTNTASATLNVAATTPFSYFTVTEWMEPSQDEARTDDNNVGTP 480
DB 421 tgmvtcmvnsvngntasatlnvaattpfsyftvcwmeppsdeartddnnvgtp 480
QY 481 VVDWETTNVTLTPQSTRTEKFTTIPVTIDNSGIPGIDVEMTKTKIIIGCFVAITLMA 540
DB 481 vvdwettntsltpqstrtektftipvtidnsgipgidevmkttkiiigcfvaitlma 540
QY 541 AVMLVIFYKMRKQHRONHHAFTTVEIINVDDTEITGTPMESHLPMPAIEHEHLNHYS 600

DB 541 avmlvifykmrkqhrqnhaptrtvelinvddteitgtpmeshlpmpaiehehlhyns 600
QY 601 YKSPFNHTTNTINSIHSSVHPEPLLRMSKDNVQETOI 640
DB 601 ykspfnhtntntinsihssvhneplllrmskdnvqetqi 640
RESULT 7
AAU00826
ID AAU00826 standard; Protein; 640 AA.
XX
AC AAU00826;
XX
DT 04-JUL-2001 (first entry)
XX
DE Human immune response protein PRO331 (UNQ292).
XX
KW Human; PRO331; UNQ292; Immune response; osteoarthritis;
KW systemic lupus erythematosus; rheumatoid arthritis; systemic sclerosis;
KW juvenile chronic arthritis; spondyloarthropathy; Sjogren's syndrome;
KW idiopathic inflammatory myopathy; polymyositis; systemic vasculitis;
KW sarcoidosis; autoimmune haemolytic anaemia; immune pancytopenia;
KW autoimmune thrombocytopaenia; idiopathic thrombocytopaenic purpura;
KW thyroiditis; Grave's disease; Hashimoto's thyroiditis;
KW diabetes mellitus; glomerulonephritis; demyelinating disease;
KW multiple sclerosis; Guillain-Barre syndrome; hepatobiliary disease;
KW chronic inflammatory demyelinating polyneuropathy; infectious hepatitis;
KW auto immune chronic active hepatitis; primary biliary cirrhosis;
KW granulomatous hepatitis; sclerosing cholangitis; ulcerative colitis;
KW inflammatory bowel disease; Crohn's disease; Whipple's disease;
KW erythema multiforme; psoriasis; asthma; allergic rhinitis; urticaria;
KW food hypersensitivity; eosinophilic pneumonia; graft rejection;
KW idiopathic pulmonary fibrosis; graft-versus-host-disease; immunogen;
KW antibody.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..44
FT /label= Signal_peptide
FT Modified-site 40..46
FT /note= "Glycine is N-myristoylated"
FT Protein 45..640
FT /label= Mature_PRO331
FT Modified-site 73..79
FT /note= "Glycine is N-myristoylated"
FT Modified-site 118..124
FT /note= "Glycine at 118 is N-myristoylated"
FT Region 183..187
FT /label= Phosphorylation_site
FT /note= "cAMP/cGMP dependent protein kinase phosphorylation site"
FT Modified-site 191..197
FT /note= "Glycine at 191 is N-myristoylated"
FT Modified-site 228..234
FT /note= "Glycine is N-myristoylated"
FT Modified-site 237..243
FT /note= "Glycine is N-myristoylated"
FT Modified-site 278..282
FT /note= "Asn is N-glycosylated"
FT Modified-site 364..368
FT /note= "Asn is N-glycosylated"
FT Modified-site 390..394
FT /note= "Asn is N-glycosylated"
FT Modified-site 391..397
FT /note= "Glycine at 391 is N-myristoylated"
FT Modified-site 412..416
FT /note= "Asn is N-glycosylated"
FT Modified-site 415..419
FT /note= "Asn is N-glycosylated"
FT Modified-site 422..428
FT /note= "Glycine is N-myristoylated"

FT Modified-site 433..439 /note= "Glycine is N-myristoylated"
 FT Modified-site 434..438 /note= "Asn is N-glycosylated"
 FT Modified-site 442..446 /note= "Asn is N-glycosylated"
 FT Modified-site 488..492 /note= "Asn is N-glycosylated"
 FT Domain 528..543 /label= Transmembrane_domain
 FT Modified-site 531..537 /note= "Glycine is N-myristoylated"
 FT Modified-site 606..610 /note= "Asn is N-glycosylated"
 PN WO200119991-A1.
 XX
 PD 22-MAR-2001.
 XX
 PF 20-MAR-2000; 2000WO-US07377.
 XX
 XX 15-SEP-1999; 99WO-US21547.
 XX (GETH) GENENTECH INC.
 PA Fong S, Goddard A, Gurney AL, Hillan KJ, Tumas D, Wood WI;
 PI WPI: 2001-226823/23.
 DR N-PSDB: AAS00162.
 DR
 XX Composition for diagnosing and treating immune related diseases, e.g.:
 FT rheumatoid arthritis and diabetes mellitus, comprises a PRO
 FT polypeptide, agonist, antagonist or fragment -
 XX
 PS Claim 31; Fig 12; 138pp; English.
 XX
 CC The sequence represents Human PRO331 (UNQ292), a protein involved in
 CC the immune response. PRO polypeptides, and (ant)agonists to them, are
 CC used in compositions for modulating infiltration of inflammatory cells
 CC into a tissue, modulating an immune response and modulating proliferation
 CC of T-lymphocytes in response to an antigen. Immune related diseases can
 CC be treated with the compositions, such as, systemic lupus erythematosus,
 CC rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis,
 CC spondyloarthropathies, systemic sclerosis, idiopathic inflammatory
 CC myopathies (e.g. polymyositis), Sjogren's syndrome, systemic vasculitis,
 CC sarcoidosis, autoimmune haemolytic anaemia (e.g. immune pancytopenia),
 CC autoimmune thrombocytopaenia (e.g. idiopathic thrombocytopaenic purpura),
 CC thyroiditis (e.g. Grave's disease, Hashimoto's thyroiditis), diabetes
 CC mellitus, immune-mediated renal disease (e.g. glomerulonephritis),
 CC demyelinating diseases of the central and peripheral nervous systems e.g.
 CC multiple sclerosis or Guillain-Barre syndrome, and chronic inflammatory
 CC demyelinating polyneuropathy, hepatobiliary diseases such as infectious
 CC hepatitis (hepatitis A, B, C, D, E and other non-hepatotropic viruses),
 CC auto immune chronic active hepatitis, primary biliary cirrhosis,
 CC granulomatous hepatitis, and sclerosing cholangitis, inflammatory bowel
 CC disease (ulcerative colitis, Crohn's disease and Whipple's disease),
 CC autoimmune or immune-mediated skin diseases (e.g. erythema
 CC multifforme or psoriasis), asthma, allergic rhinitis, urticaria,
 CC food hypersensitivity, immunologic diseases of the lung such as
 CC eosinophilic pneumonias, idiopathic pulmonary fibrosis, transplantation
 CC associated diseases including graft-versus-host-disease and graft
 CC rejection. PRO polypeptides can be used to diagnose immune related
 CC diseases, to identify inhibitors, and to stimulate the proliferation of
 CC T lymphocytes. Anti-PRO antibodies can be used to detect PRO and
 CC in diagnosis. PRO polypeptides, antibodies and (ant)agonists can be used
 CC in rational drug design.
 XX
 XX Sequence 640 AA:

Query Match 100.0%; Score 3362; DB 22; Length 640;
 Best Local Similarity 100.0%; Pred. NO. 2.3e-234;
 Matches 640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLNKTLLHPQOIMIGPRENRAFDPLLVLLALQLLVVAGLVRAQTCPSCVCSNPSKV 60
 DB 1 mlnkmtlhpqqimigprfnralfdpllvllalqlvvaglvraqtcpsvcsnpskv 60
 QY 61 ICVRKNLREVDPGISTNTRLLNLHENOQIIKVSFKHLRHLLEILQLSRNHRTIEIGAF 120
 DB 61 icvrknlrevdpgistntrllnlhengoqiiikvsfkhlrhlleilqlsrnhrtieigaf 120
 QY 121 NGLANLTLELFDNRLLTTPNGAFVYLSKLKELWLRNPNPIESIPSYAFNRPLSLRLDLG 180
 DB 121 nglanltlelfdnrllttingafvysklkelwlrnnpiesipsyafnrpslrrldlg 180
 QY 181 ELKRLSYISGAFEGLSNRLYLNLAMCNLREIPLIKLDELDSLGNHLSAIRPGSFQ 240
 DB 181 elkrlsyisgafeglsnrlrynlamcnlreiplikideidlsnghlsairpgsfq 240
 QY 241 GLMHLQKLWMTQSQIQVIERNAFDNLQSLVEINLAHNLTLLPHDLFTPLHHLERIHLLH 300
 DB 241 glmhlqklwmiqsgiqviernafdnlsqslveinlahnltllphdlftplhhlerihlhh 300
 QY 301 NPWNCNDILWLSWIKDMAPSNTACCARCNTPPNKLGRYIGELDONVETCYAPVIVEPP 360
 DB 301 npwncncdilwlswwikdmapsntaccarcntppnlkgryygeidqnyftcyapvivepp 360
 QY 361 ADLNVTEGMAELKCRASSTLSVSWITPNGTVMTGHGAYKVRIAVLSDGTINFTNVTQD 420
 DB 361 adlnvtegmaaelkcrastsitsvswitpngtvmthgaykvraviavlsdgtinfntvtqd 420
 QY 421 TGMTTCMVSNVSGNTASATLNVTAAATTPFSYFSTVTVTMEPSQDEARTDNNVGTFP 480
 DB 421 tgmtcmvsnvsgntasatlnvtaattfpysfstvtvtempsqdeartdnnvgtp 480
 QY 481 VVDWETTNVTTSLTPQSTRSEKTFPIPTDINSIGIPGIDEMVMTKIIIGCFVAITLMA 540
 DB 481 vvdwettntvttsltpqstrsektftipvtinsigipgidevmkttkiiigcfvaitlma 540
 QY 541 AVMLVIFYKMRKHRRQNHAPTRTVEINVDDEITGDTPMESHLPMPAIEHEHLNHYNS 600
 DB 541 avmlvifykmrkqhrrqnhaptrtveinvddeitgdtpmeshlmpaiehehlnhyns 600
 QY 601 YKSPFNHTTNTVNTINSIHSSVHEPLLRMSKDNVQETQI 640
 DB 601 ykspfntttntvntinsihssvhepllrmskdnvqetqi 640
 RESULT 8
 AAB80262 ID AAB80262 standard; Protein; 640 AA.
 AC AAB80262;
 XX 24-APR-2001 (first entry)
 DE Human PRO331 protein.
 KW Human; PRO; dermatological; antipsoriatic; cytostatic; antiinflammatory;
 KW antiparkinsonian nootropic; neuroprotective; vulnery; cardiac;
 KW antiangiogenic; vasotropic; antiasthmatic; antirheumatic; cancer;
 KW antiarthritic; antinfertility; antidiabetic; antiviral; diabetes;
 KW ophthalmological; gene therapy; skin disease; gastrointestinal disorder;
 KW ischaemia; inflammation.
 OS Homo sapiens.
 XX WO200104311-A1.
 XX 18-JAN-2001.
 XX 22-FEB-2000; 2000WO-US04414.
 XX 07-JUL-1999; 99US-0143048.
 XX 26-JUL-1999; 99US-0145698.

PI Ferrara N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi CJ, Gurney AL, Kljavin LJ, Napier MA, Pan J, Paoni NF;
PI Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WJ;
PI Zhang Z;
XX WPI: 2001-032160/04.
DR N-PSDB; AAF44261.
XX
XX PRO polynucleotides used to produce polypeptides used to target
PT bioactive molecules such as toxins, radiolabels or antibodies, to
PT specific cells, to cause targeted cell death -
XX
XX Claim 12; Fig 314; 935pp; English.
XX
XX The present invention describes human secreted and transmembrane PRO
CC proteins. The PRO proteins have cytostatic activity. The PRO proteins
CC can be used for targeted delivery of bioactive molecules, such as
CC toxins, radiolabels or antibodies, that cause cell death. PRO nucleotide
CC sequences, and their fragments, can be used as hybridisation probes, in
CC chromosomal and gene mapping, and in the generation of anti-sense RNA
CC and DNA. They may also be used to produce transgenic animals which are
CC used to develop and screen therapeutically useful reagents. The PRO
CC nucleotide and protein sequence can be used for tissue typing and in
CC treating cancer. Anti-PRO antibodies can be used in diagnostic assays.
CC AAF44270 to AAF44470 represent PCR primers and hybridisation probes used
CC in the isolation of human PRO sequences. AAF44087 to AAF44269 and
CC AAB65154 to AAB65300 represent human PRO polynucleotide and protein
CC sequences given in the exemplification of the present invention.
XX
XX Sequence 640 AA;
SQ

Query Match 100.0%; Score 3362; DB 22; Length 640;
Best Local Similarity 100.0%; Pred. No. 2.3e-234;
Matches 640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLNKMTHLPOQIMGPRNRPALFPLLVLLALQLLVVAGLVRAQTCPSCVCSNQFSKV 60
Db 1 mlnkmtlhpgqimgrfnralfpdlvlvllalqlvvaglvraqtcpsvcscnqfskv 60
Qy 61 ICVRKNLRVDPGISTNTRLNLENHNOIQIFKVSFKHLRHLLEILQLSRNHRITIEGAF 120
Db 61 icvrknlrrevpdgistntrlnlenhnoiqifkvsfkhrlhlleilqlsrnhritieigaf 120
Qy 121 NGLANLTLFDRNLRTTPNGAFVYLSKELWLRNPNIESIPYAFNRIPSLRRDLG 180
Db 121 nglanlntlelfdnrlrttippingafvylskelwlrnnpiesipsyafnrslrrldlg 180
Qy 181 ELKRLSYISEGAFGLSLNRLYLNLAMCNLRIPNLTPLIKDELDSLGNHLSAIRPGSFQ 240
Db 181 elkrlyisegafglslnrllynlamcnlreipnltplikdeldslgnhlsairpgsfq 240
Qy 241 GLMHQLKLMWIOSIQVTERNAFNLQSLVINLAHNLTLLPHDLFTPLHLRHLHH 300
Db 241 glmhqlkmlwiosiqvternafnqlslveinlahnltllphdlftplhlhrlhlh 300
Qy 301 NPWNCNCDILWSWIKDMAPSNFACCARCNTPPNLKGRYIGELDONFTCYAPVIVEPP 360
Db 301 npwncncdilwswikdmapsnfaccarcntppnlkgrylgldonftcyapvivepp 360
Qy 361 ADLNVTEGMAELKCRASTSLTSVTPNGVTMTHGAYKRVIAVLSGTLNFTNVTVD 420
Db 361 adlnvtegmaaelkcrastsltsvtpngvtmthgaykrviavlsgdtlnftnvtvd 420
Qy 421 TGMVTCMVNSVGNVTASATLVNVAATTPPSYFSTVVTETMEPSQDEARTDNNVGTP 480
Db 421 tgmvtcmvnsvgnvtasatlvnvaattppsyfstvtetmepsqdeartdnnvgtp 480
Qy 481 VVDWETTNTVTLTPQSTRTEKFTTIPVTDINSIGIPGIDVMTKTKIIIGCFVAITLMA 540
Db 481 vvdwettntvltlpqstrstekfttvpvtdinsigipgidvmtktkiigcfvaitlma 540
Qy 541 AVMLVIFYKRRKQHRQNHHAAPTFTVEILNVDDDEITGDTPMESHLPMPAIEHEHLNHYNS 600

Db 541 avmlvifykrmrkqhrghhapttrtvelnvdddeitgdtpmeshlpmpalehehlhyns 600
Qy 601 YKSPFNHTTTVTNTINSIHSSVHEPLLRMSKDNVQETQI 640
Db 601 ykspfnhtttvtntinsihssvhepllrmskdnvqetqi 640
RESULT 10
AAB53089
ID AAB53089 standard; Protein; 640 AA.
XX
XX AAB53089;
XX
XX 28-FEB-2001 (first entry)
XX
XX Human angiogenesis-associated protein PRO331, SEQ ID NO:137.
XX
XX Human; angiogenesis-associated protein; PRO; endothelial cell growth;
KW cardiac hypertrophy; cardiovascular disorder; endothelial disorder;
KW angiogenic disorder; atherosclerosis; osteoporosis; hypertension;
KW myocardial infarction; diabetic retinopathy; rheumatoid arthritis;
KW Crohn's disease; psoriasis; endometriosis; ulcer; wound healing; cancer;
KW Alzheimer's disease; Huntington's disease; stroke; drug screening;
XX gene therapy; transgenic animal.
XX
XX Homo sapiens.
XX
XX WO200053753-A2.
XX
XX 14-SEP-2000.
XX
XX 05-JAN-2000; 2000WO-US00219.
XX
XX 08-MAR-1999; 99WO-US05028.
XX 12-MAR-1999; 99US-0123957.
XX 14-MAY-1999; 99US-0134287.
XX 02-JUN-1999; 99WO-US12252.
XX 23-JUN-1999; 99US-0141037.
XX 20-JUL-1999; 99US-0144758.
XX 26-JUL-1999; 99US-0145698.
XX 01-SEP-1999; 99WO-US20111.
XX 08-SEP-1999; 99WO-US20594.
XX 13-SEP-1999; 99WO-US21090.
XX 15-SEP-1999; 99WO-US21547.
XX 05-OCT-1999; 99WO-US23089.
XX 30-NOV-1999; 99WO-US28313.
XX 30-NOV-1999; 99WO-US28409.
XX 02-DEC-1999; 99WO-US28564.
XX 02-DEC-1999; 99WO-US28565.
XX
XX (GETH) GENENTECH INC.
XX
XX Ashkenazi AJ, Baker KP, Ferrara N, Gerber H, Goddard A;
PI Godowski PJ, Gurney AL, Hillan KJ, Kuo SS, Mark MR, Marsters SA;
PI Paoni NF, Pitti RM, Watanabe CK, Williams PM, Wood WJ;
XX
XX WPI: 2001-090793/10.
XX N-PSDB; AAC97475.
XX
XX New isolated nucleic acid for producing a PRO polypeptide, analyzing
PT genetic disorders and treating cardiovascular, endothelial or
PT angiogenic disorders, such as atherosclerosis, wounds or cancer -
XX
XX Claim 69; Fig 52; 293pp; English.
XX
XX The invention relates to novel human angiogenesis-associated proteins
CC designated PRO proteins (AAB53064-B53097), and to nucleic acids encoding
CC PRO proteins. The invention also relates to vectors and host cells
CC comprising a PRO nucleic acid, the recombinant production of a PRO
CC protein, PRO antibodies specific for a PRO protein, fusion proteins
CC comprising a PRO protein, agonists or antagonists of a PRO protein, and
CC compounds which inhibit the expression of a PRO gene. The invention

PT Alzheimer's disease, osteoporosis, arthritis, and metabolic and liver
PT disorders comprises the 31939 polypeptide belonging to the leucine-rich
PT repeat family

XX Claim 9; Fig 1A-1C; 117pp; English.

XX The present sequence is a human leucine-rich repeat (LRR) family
CC member (referred as 31939) protein. The 31939 protein is useful for
CC treating cell proliferative or differentiative disorder, e.g., cancer
CC disorders or neuronal disorders, neurological disorders, demyelinating
CC diseases such as multiple sclerosis, degenerative diseases such as
CC Alzheimer's disease and Huntington's disease, spinocerebellar
CC degenerations, disorders of the central or peripheral nervous system,
CC bone disorders such as osteoporosis, immune disorders such as
CC rheumatoid arthritis, diabetes mellitus, cardiovascular disorders,
CC liver disorders, viral diseases, pain or metabolic disorders. They
CC are also useful as markers of disorders or disease states, or for
CC precursors or markers for the predisposition of disease states. They
CC are useful in screening and detection assays such as chromosomal
CC mapping, tissue typing, and forensic biology.

XX Sequence 713 AA;

Query Match 63.4%; Score 2131.5; DB 22; Length 713;
Best Local Similarity 61.6%; Pred. No. 1.8e-145;
Matches 409; Conservative 81; Mismatches 103; Indels 71; Gaps 9;

Qy 44 AOTCPSCVCSNQSKYICVRKRLREVPDGISTNTRLLNLHENOIQIIVKNSFKHLRHLE 103

Db 54 atscpvacscnqasqvictrrdlaevpasipvntrynlgngiqvirdtckhlrhle 113

Qy 104 ILQSRNHRTIEGAFNGLANLTLEFDNRLTTINGAFVYLSKLKELWLRNPNIESI 163

Db 114 ilqlsknlvrkievafngfngpslntelfdnrlttvqtqafeylskrlrelwnmpiesi 173

Qy 164 PSYAFNRIPSLRRLDGLKRLSYISEGAFGLSNRLRYNLAMONREIPNLPLIKLDE 223

Db 174 psyafnrpslrrldlgelkrleyiseaafeglvnlrynlmgcnlkdipnltalvrlee 233

Qy 224 LDLSGNHLSAIRPCSGFOGLHLOKWLQSQIQVIERNAFDNLQSLVEINLAHNNLFLP 283

Db 234 ldslnldlrpsfgdgltsrlklwlmhaqvatiernafodlksleelnshnlnslp 293

Qy 284 HDLFTPLHLERIHLPNPNWNCNDILWLSWIKDAPSNTPACCARNTPEPLKGRVIGE 343

Db 294 hdlftplhlrlvlnhnpncndvllswlketvpsnttccarchapaglkgrvige 353

Qy 344 LDQNYFTCYAPVIEPPADLNVTEGMAAEKLCRASTSLTSVSWITPNTGVTWTHGAYKVR 403

Db 354 ldqshftcyapvievepdtlnvtegmalekrcrtgtsmtsvnwitpntgltmthgsvrvri 413

Qy 404 AVLSDGILNFNTVNVQDTGMVTCVNSVSGNTTASATLNVTAATTP----- 450

Db 414 svldhgtlnfnvncvqdtggytcmvtsagnntasatlnsvadpvaagtgsggggpgg 473

Qy 451 -----FSYESTVVTETME--PSQD--EARTTQNN--VGPPTPVVDW-----E 485

Db 474 sgvgvgggsggytyftvtvvelctqgeaalqprgtekeppgttdgvwggrrpdaagp 533

Qy 486 TTNVNTSLTPOSTRSTKTFPIPVTDI--NSGIPGIDEVMKTKIIIGCFVAITLMAAVML 544

Db 534 asstetpaprsrptekafvtpitdvtetenalkldldvmkttkiliigcfvaitfmaavml 593

Qy 545 VIFVKMRQHRQNHQAPTRVEILNVDDET-----GDPTMESHLPM 588

Db 594 vafyklrkqghlkhqpttrveilnvdelopaasavsvaaavagvgvggdsdshlaip 653

Qy 589 AIEHEHLNHYNSYKSPFNHTTNTV-----INSIHSSVHEPFLIRMSKDNVQ 636

Db 654 alerdhlnhhyvadafkahyssnpsggcgkpgppglnsi-----hepllfksgskenvq 709

Qy 637 ETQI 640

Db 710 etqi 713

RESULT 12

AAB24073

XX ID AAB24073 standard; Protein: 653 AA.

XX AC AAB24073;

XX DT 29-JAN-2001 (first entry)

XX DE Human PRO1111 protein sequence SEQ ID NO:46.

XX KW Human; tumour; diagnosis; neoplastic disease; neoplastic cell growth;
KW proliferation; tumorigenesis; identification; cancer; cytostatic;
KW neutropic; neuroprotective; antiinflammatory; immunosuppressive;
KW immunostimulant; antiangiogenic; leukaemia; lymphoid malignancy;
KW neuronal disorder; gliial disorder; astrocytal disorder; angiogenic;
KW hypothalamic disorder; glandular disorder; macrophagal disorder;
KW epithelial disorder; stromal disorder; blastocoealic disorder;
KW inflammatory disorder; immunologic disorder.

XX OS Homo sapiens.

XX PN WO200053755-A2.

XX PD 14-SEP-2000.

XX PF 06-JAN-2000; 2000WO-US00376.

XX PR 08-MAR-1999; 99WO-US05028.

XX PR 02-JUN-1999; 99WO-US12252.

XX PR 23-JUN-1999; 99US-0141037.

XX PR 07-JUL-1999; 99US-0143048.

XX PR 26-JUL-1999; 99US-0145898.

XX PR 30-NOV-1999; 99WO-US28313.

XX PR 20-DEC-1999; 99WO-US30911.

XX PR 05-JAN-2000; 2000WO-US00219.

XX (GETH) GENENTECH INC.

XX PI Ashkenazi AJ, Baker KP, Goddard A, Gurney AL, Hillan KJ, Roy MA;
PI Watanabe CK, Wood WI;
XX WPI: 2000-572270/53.
XX N-PSDB; AAC58383.

XX PT Thirty PRO polynucleotides encoding PRO polypeptides, useful in the
XX treatment, diagnosis and prevention of cancer -

XX PS Claim 61; Fig 34; 286pp; English.

XX CC The present invention describes an isolated antibody that binds to
CC one of the human PRO proteins designated PRO12, PRO290, PRO341, PRO535,
CC PRO619, PRO717, PRO809, PRO830, PRO848, PRO943, PRO1005, PRO1009,
CC PRO1025, PRO1030, PRO1097, PRO1111, PRO1153, PRO1182, PRO1184,
CC PRO1187, PRO1281, PRO1283, PRO1317, PRO1317, PRO1710, PRO2094,
CC PRO2145 OR PRO2198. PRO antagonists can be used to inhibit tumour cell
CC growth. The PRO polypeptides and nucleotides are useful in the
CC treatment, diagnosis and prevention of cancer. The antibodies and other
CC anti-tumour compounds may be used to treat various conditions, including
CC those characterised by overexpression and/or activation of the amplified
CC PRO genes. Exemplary conditions or disorders to be treated with such
CC antibodies and other compounds include benign or malignant tumours
CC (e.g., renal, liver, kidney, bladder, breast, gastric, ovarian,
CC colorectal, prostate, pancreatic, lung, vulva, thyroid, hepatic
CC carcinomas, sarcomas, glioblastomas, and various head and neck tumours),
CC leukaemias and lymphoid malignancies, other disorders such as neuronal,
CC gliial, astrocytal, hypothalamic and other glandular, macrophagal,
CC epithelial, stromal and blastocoealic disorders, and inflammatory,
CC angiogenic and immunologic disorders, AAC58242 to AAC58366 represent PCR
CC primers and hybridisation probes used in the isolation of the human PRO

CC sequences. AAC58367 to AAC58396 and AAB24057 to AAB24089 represent human
CC PRO polynucleotide and protein sequences given in the exemplification of
CC the present invention.

PR	25-JUN-1998;	98US-0090691.	DR	WPI; 2000-072883/06.
PR	25-JUN-1998;	98US-0090694.	DR	N-PSDB; AA265033.
PR	25-JUN-1998;	98US-0090695.	XX	
PR	25-JUN-1998;	98US-0090696.	PT	Membrane-bound proteins and related nucleotide sequences
PR	26-JUN-1998;	98US-0090862.	XX	
PR	26-JUN-1998;	98US-0090863.	PS	claim 12; Fig 157; 822pp; English.
PR	01-JUL-1998;	98US-0091358.	XX	
PR	01-JUL-1998;	98US-0091360.	CC	The invention provides membrane-bound PRO polypeptides and
PR	01-JUL-1998;	98US-0091544.	CC	polynucleotides encoding them. The PRO sequences of the invention were
PR	02-JUL-1998;	98US-0091478.	CC	identified based on extracellular domain homology screening. The PRO
PR	02-JUL-1998;	98US-0091486.	CC	sequences have homology with proteins including LDL receptors, TIE
PR	02-JUL-1998;	98US-0091519.	CC	ligands and various enzymes. The membrane-bound proteins and receptor
PR	02-JUL-1998;	98US-0091626.	CC	molecules are useful as pharmaceutical and diagnostic agents. Receptor
PR	02-JUL-1998;	98US-0091628.	CC	immunoadhesins, for instance, can be used as therapeutic agents to block
PR	02-JUL-1998;	98US-0091633.	CC	receptor-ligand interactions. The membrane-bound proteins can also be
PR	02-JUL-1998;	98US-0091646.	CC	employed for screening of potential peptide or small molecule inhibitors
PR	02-JUL-1998;	98US-0091673.	CC	of the relevant receptor/ligand interaction. The PRO encoding sequences
PR	07-JUL-1998;	98US-0091978.	CC	are useful as hybridization probes, in chromosome and gene mapping and in
PR	07-JUL-1998;	98US-0091982.	CC	the generation of antisense RNA and DNA. PRO nucleic acid sequences
PR	09-JUL-1998;	98US-0092182.	CC	will also be useful for the preparation of PRO polypeptides, especially
PR	10-JUL-1998;	98US-0092472.	XX	by recombinant techniques.
PR	20-JUL-1998;	98US-0093329.	XX	
PR	30-JUL-1998;	98US-0094651.	SQ	Sequence 653 AA;
PR	04-AUG-1998;	98US-0095282.		
PR	04-AUG-1998;	98US-0095285.		
PR	04-AUG-1998;	98US-0095301.		
PR	04-AUG-1998;	98US-0095302.		
PR	04-AUG-1998;	98US-0095318.		
PR	04-AUG-1998;	98US-0095321.		
PR	04-AUG-1998;	98US-0095325.		
PR	10-AUG-1998;	98US-0095916.		
PR	10-AUG-1998;	98US-0095929.		
PR	10-AUG-1998;	98US-0096012.		
PR	11-AUG-1998;	98US-0096143.		
PR	11-AUG-1998;	98US-0096146.		
PR	12-AUG-1998;	98US-0096329.		
PR	17-AUG-1998;	98US-0096757.		
PR	17-AUG-1998;	98US-0096766.		
PR	17-AUG-1998;	98US-0096768.		
PR	17-AUG-1998;	98US-0096773.		
PR	17-AUG-1998;	98US-0096791.		
PR	17-AUG-1998;	98US-0096867.		
PR	17-AUG-1998;	98US-0096891.		
PR	17-AUG-1998;	98US-0096894.		
PR	17-AUG-1998;	98US-0096895.		
PR	17-AUG-1998;	98US-0096897.		
PR	18-AUG-1998;	98US-0096949.		
PR	18-AUG-1998;	98US-0096950.		
PR	18-AUG-1998;	98US-0096959.		
PR	18-AUG-1998;	98US-0096960.		
PR	18-AUG-1998;	98US-0097022.		
PR	19-AUG-1998;	98US-0097141.		
PR	20-AUG-1998;	98US-0097218.		
PR	24-AUG-1998;	98US-0097661.		
PR	26-AUG-1998;	98US-0097951.		
PR	26-AUG-1998;	98US-0097952.		
PR	26-AUG-1998;	98US-0097954.		
PR	26-AUG-1998;	98US-0097955.		
PR	26-AUG-1998;	98US-0097971.		
PR	26-AUG-1998;	98US-0097974.		
PR	26-AUG-1998;	98US-0097978.		
PR	26-AUG-1998;	98US-0097979.		
PR	26-AUG-1998;	98US-0097986.		
PR	26-AUG-1998;	98US-0098014.		
PR	31-AUG-1998;	98US-0098525.		
PR	16-SEP-1998;	98US-0100634.		
PR	12-JAN-1999;	99US-0115565.		
XX				
PA	(GETH) GENENTECH INC.			
XX				
PI	Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK;			
PI	Wood WL, Yuan J;			
XX				

PN WO200140466-A2.
 XX 07-JUN-2001.
 PD 01-DEC-2000; 2000WO-US32678.
 XX 01-DEC-1999; 99WO-US28301.
 PR 01-DEC-1999; 99WO-US28634.
 PR 02-DEC-1999; 99WO-US28551.
 PR 02-DEC-1999; 99WO-US28564.
 PR 02-DEC-1999; 99WO-US28565.
 PR 09-DEC-1999; 99US-0170262.
 PR 16-DEC-1999; 99WO-US30095.
 PR 20-DEC-1999; 99WO-US30911.
 PR 20-DEC-1999; 99WO-US30999.
 PR 06-JAN-2000; 99WO-US31243.
 PR 06-JAN-2000; 2000WO-US00277.
 PR 06-JAN-2000; 2000WO-US00376.
 PR 11-FEB-2000; 2000WO-US03565.
 PR 18-FEB-2000; 2000WO-US04341.
 PR 18-FEB-2000; 2000WO-US04342.
 PR 22-FEB-2000; 2000WO-US04414.
 PR 24-FEB-2000; 2000WO-US04914.
 PR 01-MAR-2000; 2000WO-US05004.
 PR 20-MAR-2000; 2000WO-US07377.
 PR 21-MAR-2000; 2000WO-US07532.
 PR 30-MAR-2000; 2000WO-US08439.
 PR 17-MAY-2000; 2000WO-US13705.
 PR 22-MAY-2000; 2000WO-US14042.
 PR 30-MAY-2000; 2000WO-US14941.
 PR 02-JUN-2000; 2000WO-US15264.
 PR 10-NOV-2000; 2000WO-US30873.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
 PI Grittisen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
 XX WPI: 2001-408281/43.
 DR N-PSDB; AAS21462.
 XX
 PT Isolated, secretory and transmembrane PRO polypeptide used to detect
 PT other PRO polypeptides, link bioactive molecules to cells expressing
 PT PRO polypeptides, and detect the presence of mammalian tumours e.g.
 PT lung, breast, prostate, cervical
 XX
 PS Claim 12; Fig 438; 813pp; English.
 XX
 CC AAU12172-AAU12446 represent novel human secretory and transmembrane
 CC PRO polypeptides. The PRO polypeptides are useful to detect other
 CC PRO polypeptides, to link bioactive molecules to cells expressing
 CC PRO polypeptides, to modulate biological activities of cells expressing
 CC PRO polypeptides, and to detect the presence of mammalian lung, colon,
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO
 CC polypeptide expression in a cell sample to that in a control sample.
 CC Some of the 275 sequences are also useful to stimulate the release of
 CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the
 CC proliferation or differentiation of chondrocytes, the proliferation or
 CC gene expression in pericyte cells, the release of proteoglycans from
 CC cartilage, the proliferation of inner ear utricular supporting cells or
 CC of T-lymphocytes, the release of a cytokine from peripheral blood
 CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
 CC to factor VIIa. The PRO polypeptides can be used in assays to identify
 CC molecules involved in binding interactions. The polynucleotides encoding
 CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,
 CC transgenic or knock out animals and can be used in gene therapy.
 XX
 SQ Sequence 653 AA;

Query Match 55.5%; Score 1865; DB 22; Length 653;
 Best Local Similarity 56.4%; Pred No. 2.8e-126;
 Matches 363; Conservative 107; Mismatches 130; Indels 44; Gaps 14;
 QY 29 VLLALQLLVVAGLVRA-----QTCPSVCSCSNOFKVICVYKRNKRLREVPDGIQSTNTRLN 82
 DB 22 vyltaqvwiilcaaaasagpqpncpsvcscsnqfsvvctriglsevpqgipnsntryln 81
 QY 83 LHENQIOIIVKNSPKHLRHLEIQLSRNHIRTIEIGAFNGLANLNTLELFDNRLLTTPNG 142
 DB 82 lmenniqmigaqdtfrhlhleviqgrnsirqlievqfnglasintlelfdnwlwtpsg 141
 QY 143 AFVYLSKLKELWNNPIESIPSAFNRIPSLRRLLDGLKRLSYISEGAFEGSLNRYL 202
 DB 142 afeylsklrelwlnnpiessipsafnrpslrmldlgelkkleyisegafegfnkyl 201
 QY 203 NLAMCNLRIPNLTPLKLDLDSGNHLSAIRPGSFGGLMHLOKLMWISQIOVIERN 262
 DB 202 nlgmcnlikdnpnltpvgleelmsgnhfpairpgsfhgisllkklwvmsqvalier 261
 QY 263 FDNLOSLVEINLAHNNLTLLPHDLFTPLHLERLHLHNPNCNCIDLILWSWIKDMAPS 322
 DB 262 fdglasveinlahnnlssphdlftplrylvelhlhnpwncdcdilwlawlreyipt 321
 QY 323 NTACCARCNTPPNLKGRYIGELDQNYFTCYAPVIVEPPADLNVTEGMAALCKRASTSLT 382
 DB 322 nstccgrchapmbmrgrylvevdqasfcsapfimdaprdinisegrmaelker-tpgms 380
 QY 383 SVSWITPNTGVTHGAYKVIKRIAVLSODTGLNFTVNTVODTGMTCWMSVNGTNTASATLN 442
 DB 381 svkwlplngtvlshashrprisvldgtlnfshvllsdgtvytcmvtnvagnsnasayln 440
 QY 443 VTAA--TTTPFSYFSTVTVTETMBPSQDEARTDNNVGPVVDWETNV-----TTSLT 494
 DB 441 vstaelnstnyfctvtvvetelsped---trkykvpv-----tstgyqpsaystst 492
 QY 495 P--QSTRSTKTEFTIPVTDINSIGI--PGIDEVMTKTKIIGCFVAITLMAAVMLVIFYKMR 551
 DB 493 vliqttr-vpkqvavpatdtdtkmqtsldevmkttkiigcfvavtllaaamlivfyklr 551
 QY 552 KQHRQNHHPARTVEIINVDDEITGDTPM-----ESHLPMPTAHEHNLHNS 600
 DB 552 krhqrstvttaartveilqvdvdiipaatsaaataapsgvsgegavvptl-hdhin-ynt 609
 QY 601 YKSPFNHTTNTI--NSIH---SSVHPPLITRMNSKDNVOETQI 640
 DB 610 ykpahgahwtenslgnshlptvtvtisepylqthtkkvtqql 653

Search completed: August 26, 2002, 15:40:58

Job time: 206 sec

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	1376.5	40.9	421	2	T46266	hypothetical prote
2	417.5	12.4	1091	2	A58532	glial cell membran
3	410.5	12.2	361	2	A53860	chondroadherin pre
4	406	12.1	707	2	JC7763	neuroal leucine-r
5	399	11.9	1531	2	T42218	slit-1 protein hom
6	374.5	11.1	1523	2	T13953	MEGF5 protein - ra
7	350	10.4	1469	2	B36665	slit protein 2 pre
8	350	10.4	1480	2	A36665	slit protein 1 pre
9	348.5	10.4	789	2	T28714	hypothetical prote
10	348.5	10.4	1355	2	T28715	hypothetical prote
11	344.5	10.2	1025	2	T42626	secreted leucine-r
12	337.5	10.0	1535	2	S46224	peroxidase - frui
13	335	10.0	605	2	JC5939	insulin-like growt
14	329.5	9.8	560	2	A60164	platelet membrane
15	320	9.5	605	2	A41915	insulin-like growt
16	312	9.3	603	2	JC1282	insulin-like growt
17	311	9.3	603	2	JC6128	insulin-like growt
18	302	9.0	359	1	NBHUC8	deccorin precursor
19	296.5	8.8	594	2	T23841	hypothetical prote
20	296.5	8.8	626	1	NBHUIA	platelet glycoprot
21	293	8.7	357	2	S24317	deccorin precursor
22	293	8.7	360	2	S06280	deccorin precursor
23	293	8.7	360	2	I47020	deccorin - rabbit
24	290	8.6	354	2	A55454	deccorin precursor
25	289	8.6	907	2	JG0193	G protein-coupled
26	287.5	8.6	536	2	A34901	lysine carboxypept
27	287.5	8.6	610	2	T23036	hypothetical prote
28	287	8.5	354	2	S29145	deccorin precursor
29	286	8.5	420	2	A53531	oncofetal trophobl

A:Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs
A:Reference number: Z14126; MUID:98360089

A:Accession: T13953

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1523 <NAK>

A:Cross-references: EMBL:AB011531; NID:g3449291; PIDN:BAA32461.1; PID:g3449292

C:Genetics:

A:Gene: MEGF5

Query Match 11.1%; Score 374.5; DB 2; Length 1523;
Best Local Similarity 25.4%; Pred. No. 2.2e-16;
Matches 120; Conservative 49; Mismatches 147; Indels 157; Gaps 9;

QY 31 LALQLLVAGLVRAQTCPSCVCSNQSKVICVRKNLREVPDGIISTWTRLLNHNQIQI 90

DB 18 LALALSLSGPPAAACPTKTCSS--AASVDCHGLGLRAVRGPRIPAERLDLDORNNITR 75

QY 91 IKVNSFKHLRHLLEILQLSRNHRTIEIGAFNGLANLTLELFDNRLTIPNGAFVYLSKL 150

DB 76 ITKWDTGLKKNLVLHLEDNQSVIERGAFODLAKQLERLRLNKNKQLVLPPELLFQSTPKL 135

QY 151 KELWLNPNISISYAFNRTPSRRDLDBELKRLSYISEGAFGLSNLRYNL----- 204

DB 136 TRLDLSENIQIOPKAPRGVGTGVKNLQL-DNNHISCIEDGAFRALRDLLEITLNNNNIS 194

QY 205 ----- 204

DB 195 RILVTSFNHMPKIRTLRLSHNHLVCDCHLAWLSDWLRQRRITGFTLCMAPVHLRGFSA 254

QY 205 -----AMCN-----LREIP-NLTPLI----- 219

DB 255 DVQKKEYVCPGHPSEAPACNANSLSCPSACSNIVDCRGKGLTEIPANLPEGIVEIRL 314

QY 220 -----KLDELDSGNHLSAIRPGSQ----- 240

DB 315 BONSISIPAGAFQYQKLRIDISKNOISDIAPDAFQGLKSLTSLVLYGNKITEIPKGL 374

QY 241 --GLMHLQKLMWIOSQIOVIERNADFNLSLVEINLAHNNTLLPHDLFTPLHLHERLHL 298

DB 375 FDGLVSLQLLLNANKINCLRVNTFODLONLNLSSLDYDNKLOTISKGLFAPLQSIQTILH 434

QY 299 HHNPWNCNDILWLSWTKMAPNTACCACNTPPNLKGRYIGELDONQYTC 351

DB 435 AQNPFCVCDHLKWLADYLDQ-NPIETS-GARCSSPRLANKRKRIQSKKPRC 485

RESULT 7

B36665

C:Species: Drosophila melanogaster

C:Date: 30-Apr-1991 #sequence_revision 30-Apr-1991 #text_change 19-May-2000

C:Accession: B36665

R:Rothberg, J.M.; Jacobs, J.R.; Goodman, C.S.; Artavanis-Tsakonas, S.

Genes Dev. 4, 2169-2187, 1990

A:Title: slit: an extracellular protein necessary for development of midline glia and cc

A:Reference number: A36665; MUID:91099665

A:Accession: B36665

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1469 <ROT>

A:Cross-references: GB:X53959

C:Genetics:

A:Gene: FlyBase:slit

C:Cross-references: FlyBase:FBgn0003425

C:Superfamily: unassigned EGF-related proteins; EGF homology; leucine-rich alpha-2-glyco

F:66-91/Domain: proteoglycan amino-terminal homology <PAH2>

F:101-124/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>

F:125-148/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>

F:149-172/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>

F:173-196/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>

F:197-220/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>

F:228-272/Domain: proteoglycan carboxyl-terminal homology <PCS1>

F:288-313/Domain: proteoglycan amino-terminal homology <PAH2>

F:323-346/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>

F:347-370/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>

F:371-394/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>

F:395-418/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>

F:419-442/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>

F:450-494/Domain: proteoglycan carboxyl-terminal homology <PCS2>

F:512-537/Domain: proteoglycan amino-terminal homology <PAH3>

F:547-571/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR11>

F:572-595/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR12>

F:596-619/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR13>

F:620-643/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR14>

F:651-695/Domain: proteoglycan carboxyl-terminal homology <PCS3>

F:708-733/Domain: proteoglycan amino-terminal homology <PAH4>

F:743-766/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR15>

F:767-790/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR16>

F:846-890/Domain: proteoglycan carboxyl-terminal homology <PCS4>

F:1028-1061/Domain: EGF homology <EGF2>

F:1068-1093/Domain: EGF homology <EGF2>

F:1115-1148/Domain: EGF homology <EGF1>

Query Match 10.4%; Score 350; DB 2; Length 1469;

Best Local Similarity 23.9%; Pred. No. 8.4e-15;

Matches 109; Conservative 65; Mismatches 152; Indels 130; Gaps 9;

QY 46 TPCSVCSNQSKVICVRKNLREVPDGIISTWTRLLNHNQIQIIVNSFKHLRHLLEIL 105

DB 294 SCPHPCADGI--VDCREKSLTSVPVTLPDPTDVRLEQNFITELPKPSSSFRRLRI 351

QY 106 QLSRNHIRTIEIGAFNGLANLTLELFDNRLTIPNGAFVYLSKLKELWLNPNIESIPS 165

DB 352 DLSNNISRIADHSLGKQLTTLVLYGNKIKDLPSPGVFKGLSLRLLLNANEISCIRK 411

QY 166 YAFNRISRLRDLGELKRLSYISEGAFGLSNLRYNLAM-----CNLR----- 210

DB 412 DAFRDLHSLSLSLYD--NNIOSLANGTFDAMKSMKTVHLAKNPFITCDNLRNLADYLHN 470

QY 211 ----- 210

DB 471 PIETSGARCESPKRMHRRRIEBSLEEKFKCSWGLRMKLSGECRMDSDCPAMCHCEGTV 530

QY 211 -----EIPNLTP-----TKLD-----ELDLSGNHLSAIRPGS 238

DB 531 DCTGRLKEIPRIDPLHTTELLNDNELGRSSDGLFGRLPHLVKLELKRNLQLTGEPNA 590

QY 239 FQGLMHLQKLMWIOSQIOVIERNADFNLSLVEINLAHNNTLLPHDLFTPLHLHERLHL 298

DB 591 FEGASHIQELQIENKIKESNKMPLGLHQLTLLNLDNQISCVMPGSEFHLNLSLNL 650

QY 299 HHNPWNCNDILWLSWTKMAPNTACCACNTPPNLKGRYIGELDONQYTCYAPVIVE 358

DB 651 ASNPFNCNCHLAWAEVCVRKKSNGGA--ARCGAPSKVRDQVQIKDLPHSEFKCSS----- 703

QY 359 PPADLVNTEGMAELKCRASLTSLTSVSWITPNGVTM 394

DB 704 -----ENSEGLGDCGCPSPCTCT-----GTVV 726

RESULT 8

A36665

C:Species: Drosophila melanogaster

C:Date: 30-Apr-1991 #sequence_revision 30-Apr-1991 #text_change 17-Nov-2000

C:Accession: A36665; A31640; S13523

R:Rothberg, J.M.; Jacobs, J.R.; Goodman, C.S.; Artavanis-Tsakonas, S.

Genes Dev. 4, 2169-2187, 1990

A:Title: slit: an extracellular protein necessary for development of midline glia and

A:Reference number: A36665; MUID:91099665

A:Accession: A36665

A:Status: preliminary

A:Molecule type: mRNA

RESULT 10

T28715
hypothetical protein T21D12.9b - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C:Accession: T28715
R:Woessner, J.
submitted to the EMBL data library, August 1997
A:Description: The sequence of C. elegans cosmid T21D12.
A:Reference number: Z20514
A:Accession: T28715
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1355 <WOE>
A:Cross-references: EMBL:AF016687; PIDN:AAC48095.1; GSPDB:GN00022; CESP:T21D12.9b
A:Experimental source: strain Bristol N2; clone T21D12
C:Genetics:
A:Gene: CESP:T21D12.9b
A:Map position: 4
A:Introns: 38/2; 84/2; 132/2; 204/2; 275/2; 351/2; 519/3; 615/1; 714/1; 758/2; 786/3; 84

Query Match 10.4%; Score 348.5; DB 2; Length 1355;
Best Local Similarity 24.3%; Pred. No. 9.5e-15;
Matches 131; Conservative 85; Mismatches 233; Indels 89; Gaps 19;
QY 75 STNTRL-LNLHENQIQIKVNSFKHLRHLLETLQLSRNHRTIEIGAFNGLANLNTLELFLQ133
Db 101 SFNTLVTLKLNHITLNFQSFRLKRLKLESLDTRNMIREVRLAFNQLPSPQNVSLAR 160
QY 134 NRTLTTPNGAFVYLSKLEKWLNRNPIESIP-----SYAFNRI----- 171
Db 161 NDVYRLDDGEYACEGLKHLNLSNRVQAVTEGWMFGLSLEVLDSYNQIOSPHISSWS 220
QY 172 --PSLRDLGELKRLSYISEGAFEGSLNRYLNLMCNLRIPN--LTPLIKLDELDSL 227
Db 221 HTPKRLKWSL-HSNRIQSLPSGSPRVLRQLEELTILSANSIDSLHKFALVGMSSLHKLDLS 279
QY 228 GNHLS-AIRPGSF--OGLMHLQKLMWIOSQIOVIERNAPDNLSLVEINLAHNNLTLP 283
Db 280 SNTLVACVEDGAVLYNTSMFLRLSRTNNQLRVIPRAFERFFALELOLTDNPIATIH 339
QY 284 HDLFTPLHLERLHLHNPNCNCDILWSWIKDMAPSNTACCARGTTPNKLGRYIGE 343
Db 340 PEAPEPL-ELKRLVMSNSSLCDQISWLASVIYRLKDKSIIIAKCSYPPPLADLYVA 398
QY 344 LDQNYFTCY----APVIEPPADLVNTEGMAELKRA-STSLTSVSW----- 386
Db 399 IDTANLCHNDSPRAKIVRQPEVSTLIGEKAREFTCNVYGASPLSIEWRMENQOPRVLV 458
QY 387 -----ITPNCVTNTHGAYKVRIVASLDGTLNFTNVTVDQTMVTCMVSNSVGNNTASAT 440
Db 459 QDSATFSLNKTAVVNTGTFDER--ELAAEALLDNVAMTNSQVQVARNRFG--SDFSH 515
QY 441 LNVTAATTTTFSYSTVIVETM-----EPSQDEARTDNNVGPVVDWETNVTTSITLP 495
Db 516 VKLVQVQAPKFTY---TPEDMPLLVQGTAKFLCAATGT---PRPEIKWAFQI-----P 563
QY 496 QSTRSTKFTTIPVDINSIGIPGIDVWTKTIIGCF-----VAITLMAAVMLVIF 547
Db 564 FPAAEARRLVYTPND-----HIYIMNVTKEGQAYTCHATNVAGQTQASANLIVF 614

RESULT 11

T42626
secreted leucine-rich repeat-containing protein SLIT2 - mouse (fragment)
N:Alternate names: neurogenic extracellular slit protein
C:Species: Mus musculus (house mouse)
C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-Jan-2000
C:Accession: T42626
R:Holmes, G.P.; Negus, K.; Burridge, L.; Raman, S.; Algar, E.; Yamada, T.; Little, M.H.
Mech. Dev. 79, 57-72, 1998

A:Title: Distinct but overlapping expression patterns of two vertebrate slit homologs
A:Reference number: Z22177; MUID:99279238
A:Accession: T42626
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1025 <HOL>
A:Cross-references: EMBL:AF074960; NID:g4151258; PID:g4151259; PIDN:AAD04345.1
C:Genetics:
A:Gene: Slit2

Query Match 10.2%; Score 344.5; DB 2; Length 1025;
Best Local Similarity 24.3%; Pred. No. 1.2e-14;
Matches 102; Conservative 62; Mismatches 149; Indels 107; Gaps 8;
QY 47 CPSVCSNCSNFSKVICVRKNLRVDPDGISINTNRLNLHENQIQIKVNS-FKHLRHLLEIL 105
Db 2 CPEKCRCEG--TTVDCSNQRLNKIPDHIPQYTAEELNNNEFTLEATGIFKKLPQURXI 59
QY 106 QLSRNHRTIEIGAFNGLANLNTLELFDNRLTTPNGAFVYLSKLEKWLNRNPIESIPS 165
Db 60 NFSNNKITDIEGAFEGASGVNEILLTSNRLNENYQHKMGLESKLTLMLSNRISCVGN 119
QY 166 YAFNRIPSLRRLDGLKRLSYISEGAFEGSLNRYLNLM-----CN----- 208
Db 120 DSFTGLGSVRLSLYD-NQITTVAPGAFDXLHLSLSTLNLNLANPENCNCHLAWLGEWLRRK 178
QY 209 -----LRPIP-----NLTPLIK----- 220
Db 179 RIVTGNPRCKPKYFLKEIPIQDVAIQDFTCDGNDNDNSCPLSRCPSECTCLDTXVRCSN 238
QY 221 -----LDELDSGNHLSAIRPGSFQGLMH 244
Db 239 KGLKVLPGKIPKDVTELYLDGNQFTLVLPKELSNYKHLTLIDLNNRISTLSNQAFSNMTQ 298
QY 245 LQKLMWIOSQIOVIERNAPDNLSLVEINLAHNNLTLLPHDLFTPLHLERLHLHNPWN 304
Db 299 LLTILSYNRLCIPPTFTDGLKSLRLSLHSGNDISVVPEGAFNDLSALSHLAIGANPLY 358
QY 305 CNCIDILWSWIKDMAPSNTACCARGTTPNKLGRYIGELDNQNYFTCYAPVIEPPADLN 364
Db 359 CDCNNQWLSDMVK--SEYKEPGIARCAAGGEMADKLLTTPSKKFTCOGPMDDITIQACN 416

RESULT 12

S46224
peroxidase - fruit fly (Drosophila sp.)
C:Species: Drosophila sp.
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Oct-2000
C:Accession: S46224
R:Nelson, R.E.; Fessler, L.I.; Takagi, Y.; Blumberg, B.; Keene, D.R.; Olson, P.F.; Pa
EMBO J. 13, 3438-3447, 1994
A:Title: Peroxidase: a novel enzyme-matrix protein of Drosophila development.
A:Reference number: S46224; MUID:94341255
A:Accession: S46224
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1535 <NEU>
A:Cross-references: GB:U11052; NID:g531384; PID:AAA61568.1; PID:g531385
C:Superfamily: peroxidase; myeloperoxidase homology; proteoglycan amino-terminal hom
F:19-44/Domain: proteoglycan amino-terminal homology <PAH4>
F:661-1350/Domain: myeloperoxidase homology <MPX>

Query Match 10.0%; Score 337.5; DB 2; Length 1535;
Best Local Similarity 25.6%; Pred. No. 5.9e-14;
Matches 106; Conservative 58; Mismatches 141; Indels 109; Gaps 13;
QY 26 LLYVLLALQLLVAGLVRAOTCPSCVCSNCSNFSKVICVRKNLRVDPDGISINTNRLNLHE 85
Db 6 LMLQLLGL-LLLLAGGVQSVYCPAGCTCLER--TVCRIRAKLSAVP----- 48
QY 86 NQIOIKVNSFKHLRHLLETLQLSRNHRTIEIGAFNGLANLNTLELFDNRLTTPNGAFV 145

Db 49 -----KLPDQTTLDRFNHTEELPANAFSGLAQLTTLFLNDNLAYLQDGN 97
QY 146 YLSKLKELWLNRPISPYAFNRIPSLRRLDGLKRLSYISEGAFEGSLNRYLNLA 205
Db 98 GLTALREVYLNNNKSLRPATIFQRMPLR-----EGIF--LEN----- 133
QY 206 MCNLRIPNLTPLTKDELDSGNHLSAIRPGSPQGLMHQLKLMQISQIQVIERNAFDN 265
Db 134 -----NDIQLPAGL-----FDN 146
QY 266 LOSLVEINLAHNNTLLPHDLFTPLHLERHLHNN--PWNCRNDI--LMSLWIKMAP 321
Db 147 LPRNLRLMYNNKLTQLPVDGFNRLNNKRLDGNDAIDCNCGVYSLWRRWHL-DVQR 205
QY 322 SNTACCACRCPNPLKGRYIGELDONFTCYAPVIVPEPPADLNVTGMAAEKCRASL 381
Db 206 QLVNSITCAAPQMLQNGQSSGEHFKCAKQFIVAPQDAQVAGEQVELSECV-TGL 264
QY 382 --TSVSWITPNTGTHGAYKVRIVLSDGTNLNFTVYQDTGMTCMVNSVSG 433
Db 265 HRPOITWN--HNTQELGLEQTOAEILPSSGSLHRSADTSDMGIIQCIARNEMG 316
RESULT 13
JC5239
insulin-like growth factor acid-labile chain - baboon
C:Species: Papio sp. (baboon)
C:Date: 17-Apr-1997 #sequence_revision 09-May-1997 #text_change 09-May-1997
C:Accession: JC5239
R:Dehanty, P.; Baxter, R.C.
Biochem. Biophys. Res. Commun. 227, 897-902, 1996
A:Title: The cloning and expression of the baboon acid-labile subunit of the insulin-like
A:Reference number: JC5239; MUID:97040714
A:Contents: liver
A:Accession: JC5239
A:Molecule type: mRNA
A:Residues: 1-605
C:Comment: This factor is structurally related to proinsulin and have insuline-like meta

Query Match 10.08; Score 335; DB 2; Length 605;
Best Local Similarity 20.88; Pred. No. 2.5e-14;
Matches 125; Conservative 64; Mismatches 125; Indels 286; Gaps 13;
QY 31 LALQLLYVA-----GLVRAQTCPSCVSCS-----NQFSKVICVRKNLR 68
Db 8 LALALLLSWALGPRSLGAEPTGCEAECPCATCACSYDDEVNELS-VFCSSRLT 66
QY 69 EVPDGISTNRL----- 80
Db 67 RLDPGIPGQTQALWLDNNLSIPPAFRNLSSLAFLNLQGLGSLPEQALLGLENLCH 126
QY 81 LNLHENOIQIKVNSFKHLRHLLEIQLSRNHRTIEIGAFNLANLNTLEFDNRLTIP 140
Db 127 LHLERNQLSLAVGTFAITPALALLGSLNNKSLRDLGLFEGGLNMDNLGNWSLAVIP 186
QY 141 NCAG-----VY-----LSKLKELWLNRPISPYAFNRIPSLR 176
Db 187 DAAFRGLGLRELVLGNRLAYLOPALFSLGAELELDLSNLRANKANVFAQLPRLOK 246
QY 177 L----- 177
Db 247 LYLDNRLIAAVPAGFLGLKALRWLDLSHNRVAGLLEDTFPGLLGLRLVRLSHNAIASLR 306
QY 178 -----DLGELK-----RLSYISEGAFEGLSNRY 201
Db 307 PRTEDLHLELOLGNHNRIRQLAERSFEGLGQLEVLTLDRNQEQYKVGAGFLGTNAV 366
QY 202 LNLAMCNLRIPN--LTPLIKDELDSGNHLSAIRPGSPQGLMHQLKLMQISQIQVIE 259
Db 367 MNLGNCNLRNLPQVFRGLGKLSHLHLEGSLGIRPHPTFAGLSGLRLFLKONGLVGIE 426

QY 260 RNAPDNQLQSLVEINLAHNNTLLPHDLFTPL----- 290
Db 427 EOSLWGLAQLLELDLTSNQLTHLPHQLFQGLKLEYLLSHNRLAELPADALGPLORAFW 486
QY 291 -----HH-----LERIHLHHPWNCNCDILW 311
Db 487 LDYSHNRLAELPGSLASLGLRLRYLNLNNSLRTFTTTPQFGLERLWLEGNPWDCSCLPKA 546
QY 312 LSNWIKMAPSNTACCAR-----CNTPPNLKGRYIGELDONFY 349
Db 547 L-----RDFALQNFSVPRFVQATCEGDDCQPPVYNNITCASPEVAGLDRLDLGEAHF 602
RESULT 14
A60164
platelet membrane glycoprotein V precursor - human
C:Species: Homo sapiens (man)
C:Date: 12-Jan-1993 #sequence_revision 24-Feb-1994 #text_change 05-Nov-1999
C:Accession: A48030; A60164; A35483; B35483; A60432; A47507; S34329
R:Lanza, F.; Morales, M.; de La Salle, C.; Cazenave, J.P.; Clemetson, K.J.; Shimomura
J. Biol. Chem. 268, 20801-20807, 1993
A:Title: Cloning and characterization of the gene encoding the human platelet glycoprotein
A:Reference number: A48030; MUID:94012616
A:Accession: A48030
A:Molecule type: DNA
A:Residues: 1-560 <LA2>
A:Cross-references: EMBL:Z23091; NID:G312501; PIDN:CAA80637.1; PID:G312502
R:Shimomura, T.; Fujimura, K.; Maehama, S.; Takemoto, M.; Oda, K.; Fujimoto, T.; Oyama
Blood 75, 2349-2356 1990
A:Title: Rapid purification and characterization of human platelet glycoprotein V: th
A:Reference number: A60164; MUID:90275263
A:Accession: A60164
A:Molecule type: protein
A:Residues: 365-384 'X', 386-390 'X', 392-395 'X', 397-188-208 'I', 210; 27-50 'X', 52-53 'I',
'XX', 108 'T', 61-72 'TK', 75-77 'V', 56-57 'G', 479-487 'X', 489-498 'X', 500 'X', 502-503,
R:Roth, G.J.; Church, T.A.; McMullen, B.A.; Williams, S.A.
Biochem. Biophys. Res. Commun. 170, 153-161, 1990
A:Title: Human platelet glycoprotein V: a surface leucine-rich glycoprotein related t
A:Reference number: A35483; MUID:90321220
A:Accession: A35483
A:Molecule type: protein
A:Residues: 145-166 'I', 168-169 'X', 171-172 <ROT>
A:Note: this proteolytic fragment was designated peptide M392
A:Accession: B35483
A:Molecule type: protein
A:Residues: 121-129 'W', 131-135; 466-468 'X', 470 <RO2>
A:Note: this material was designated peptide M393 but may contain two peptides
A:Accession: C35483
A:Molecule type: protein
A:Residues: 252-266 'H', 268-272 'X', 274-279 'I', 281-284 'I', 286 <RO3>
A:Note: this proteolytic fragment was designated peptide M401
Thromb. Res. 53, 31-44, 1989
A:Title: Platelet membrane glycoprotein V: characterization of the thrombin-sensitive
A:Reference number: A60432; MUID:89162331
A:Accession: A60432
A:Molecule type: protein
A:Residues: 477-478 'FX', 481-485 'E', 487 'V', 489-492 'NQ', 495 'E', 497-498 <ZAF>
R:Hickey, M.J.; Hagen, F.S.; Vagi, M.; Roth, G.J.
Proc. Natl. Acad. Sci. U.S.A. 90, 8327-8331, 1993
A:Title: Human platelet glycoprotein V: characterization of the polypeptide and the r
A:Reference number: A47507; MUID:93391348
A:Accession: A47507
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-560 <RES>
A:Cross-references: GB:L11238; NID:G388759; PIDN:AAA03069.1; PID:G388760
C:Comment: This platelet membrane protein is a substrate for thrombin.
C:Comment: The amino end of the intact protein is blocked.
C:Comment: This protein is absent in Bernard-Soulier syndrome.
C:Genetics:
A:Gene: GDB:GP5
A:Cross-references: GDB:230236; OMIM:173511

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 15:41:28 ; Search time 13.57 Seconds
(without alignments)
1826.125 Million cell updates/sec

Title: US-09-905-056-292
Perfect score: 3362
Sequence: 1 MLNRMTHLPHQIMIGPRFNR.....VHEPLLRMSKDNVQETQI 640

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	438.5	13.0	713	1 GAC1_HUMAN	O75325 homo sapien
2	411.5	12.2	361	1 CHAD_BOVIN	Q27972 bos taurus
3	369.5	11.0	567	1 GPV_MOUSE	O08742 mus musculus
4	361.5	10.8	567	1 GPV_RAT	O08770 rattus norv
5	350	10.4	1480	1 SLIT_DROME	P24014 drosophila
6	335	10.0	605	1 ALS_PAPHA	O02833 papio hamad
7	329.5	9.8	560	1 GPV_HUMAN	P40197 homo sapien
8	320	9.5	605	1 ALS_HUMAN	P35858 homo sapien
9	312	9.3	603	1 ALS_RAT	P35859 rattus norv
10	311	9.3	603	1 ALS_MOUSE	P70389 mus musculus
11	302	9.0	359	1 PGS2_HUMAN	P07585 homo sapien
12	299	8.9	360	1 PGS2_HORSE	O46542 equus caball
13	296.5	8.8	626	1 GPBA_HUMAN	P07359 homo sapien
14	295	8.8	360	1 PGS2_CANFA	Q29393 canis fami
15	295	8.8	360	1 PGS2_PIG	Q9XSD9 sus scrofa
16	293	8.7	357	1 PGS2_CHICK	P28675 gallus gall
17	293	8.7	360	1 PGS2_BOVIN	P21793 bos taurus
18	293	8.7	360	1 PGS2_RABIT	Q28888 oryctolagus
19	290	8.6	354	1 PGS2_MOUSE	P28654 mus musculus
20	287.5	8.5	536	1 CBP8_HUMAN	P22792 homo sapien
21	287	8.5	354	1 PGS2_RAT	Q01129 rattus norv
22	276	8.2	353	1 KERA_COTJA	Q9DE66 coturnix co
23	276	8.2	966	1 Y918_HUMAN	O94991 homo sapien
24	274.5	8.2	331	1 PLIB_AGKBL	O93233 agkistrodon
25	274	8.1	353	1 KERA_CHICK	O42235 gallus gall
26	271.5	8.1	368	1 PGS1_HUMAN	P21810 homo sapien
27	271.5	8.1	369	1 PGS1_CANFA	O02678 canis fami
28	271.5	8.1	372	1 PGS1_HORSE	O46403 equus caball
29	270.5	8.0	369	1 PGS1_MOUSE	P28653 mus musculus
30	270.5	8.0	369	1 PGS1_RAT	P47853 rattus norv
31	269.5	8.0	369	1 PGS1_SHEEP	O46390 ovis aries
32	268.5	8.0	369	1 PGS1_BOVIN	P21809 bos taurus
33	268	8.0	423	1 OMD_RAT	Q9Z1S7 rattus norv

ALIGNMENTS

RESULT 1	GAC1_HUMAN	STANDARD;	PRT;	713 AA.
ID	GAC1_HUMAN	STANDARD;	PRT;	713 AA.
AC	O75325;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Glioma amplified on chromosome 1 protein precursor.			
GN	GAC1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Glial tumor;			
RX	MEDLINE=98324709; PubMed=9662332;			
RA	Malfoy B., Almeida A., Zhu X.X., Vogt N., Tyagi R., Muleris M.,			
RA	Dutrillaux A.-M., Dutrillaux B., Ross D., Hanash S.;			
RT	"GAC1, a new member of the leucine-rich repeat superfamily on			
RT	chromosome band 1q32.1, is amplified and overexpressed in malignant			
RT	gliomas."			
RL	Oncogene 16:2997-3002(1998).			
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).			
CC	-1- TISSUE SPECIFICITY: OVERAMPLIFIED IN MALIGNANT GLIOMAS.			
CC	-1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.			
CC	-1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.			
CC	-1- SIMILARITY: CONTAINS 11 LEUCINE-RICH REPEATS (LRR).			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AF030435; AAC39792.1;			
DR	MM; 605492;			
DR	InterPro; IPR003598; Iq_MHC.			
DR	InterPro; IPR003598; Iq_c2.			
DR	InterPro; IPR001611; LRR.			
DR	InterPro; IPR000483; LRR_Cterm.			
DR	InterPro; IPR000372; LRR_Nterm.			
DR	InterPro; IPR003592; LRR_out.			
DR	InterPro; IPR003591; LRR_typ.			
DR	Pfam; PF00047; Iq; 1.			
DR	Pfam; PF00560; LRR; 10.			
DR	Pfam; PF01463; LRRCT; 1.			
DR	Pfam; PF01462; LRRNT; 1.			
DR	PRINTS; PR00019; LEURICHRPT.			
DR	SMART; SM00408; IGC2; 1.			
DR	SMART; SM00370; LRR; 6.			
DR	SMART; SM00082; LRRCT; 1.			
DR	SMART; SM00013; LRRNT; 1.			
DR	SMART; SM00369; LRR_TYP; 2.			

P82963 tribolium c
O99mb1 mus musculus
P12024 drosophila
O9nyk1 homo sapien
O94933 homo sapien
P51888 homo sapien
O15455 homo sapien
P58661 mus musculus
O62702 bos taurus
O35103 mus musculus
O60938 homo sapien
P23515 homo sapien

KW Immunoglobulin domain; Transmembrane; Glycoprotein; Repeat;
KW Leucine-rich repeat; Signal.

FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 713 GLIOMA AMPLIFIED ON CHROMOSOME 1 PROTEIN.
FT DOMAIN 19 630 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 631 651 POTENTIAL.
FT DOMAIN 652 713 CYTOPLASMIC (POTENTIAL).
FT REPEAT 92 115 LRR 1.
FT REPEAT 116 139 LRR 2.
FT REPEAT 140 163 LRR 3.
FT REPEAT 165 187 LRR 4.
FT REPEAT 188 211 LRR 5.
FT REPEAT 213 235 LRR 6.
FT REPEAT 236 259 LRR 7.
FT REPEAT 261 283 LRR 8.
FT REPEAT 309 333 LRR 9.
FT REPEAT 334 357 LRR 10.
FT REPEAT 359 385 LRR 11.
FT DOMAIN 438 504 IG-LIKE C2-TYPE DOMAIN.
FT DISULFID 445 497 BY SIMILARITY.
FT CARBOHYD 94 94 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 381 381 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 555 555 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 583 583 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 713 AA; 78798 MW; EC8BC0FD240C9396 CRC64;

Query Match 13.0%; Score 438.5; DB 1; Length 713;
Best Local Similarity 25.2%; Pred. No. 2.3e-22;
Matches 147; Conservative 86; Mismatches 228; Indels 123; Gaps 12;

QY 20 RALFDPLVLLALQLVAVLRAQTCPSVCSG-----NQFSKVICVRNLR 68
DB 2 RLLVAPLLAVACATAAPVPHVPCPCQACQIRPWTPRSSYREATTVDNCLFLT 61
QY 69 EYPDGIPTNTPLLNHENOQIIKVNSEFKHLRLEILQLSRNRHTTEIGAFNGLNLT 128
DB 62 AVPPALPAGTQTLQLQNSIVRVDSGLYLANTELDSQNSFSDARDCDFHALPQLLS 121
QY 129 LEFDNRLTTPNGAFVYLSKELWLRNPNIESIPYAFNRIPSLRLDLGE----- 181
DB 122 LHLBENQLTREDSHFSAGLASQLBYLNHNQLYRIAPAFSGLSNLLRLHLSNLLRAID 181
QY 182 -----LKRSLYISEGA-----FEGLSNRLYLNAMCNIREIPN----- 214
DB 182 SRWFEMLPNLEILMIGGNKVDAILDMNFRPLANERSLVLAGMNLREISDYALEGLQSLES 241
QY 215 -----LPLIKLDELDSGNHLSAIRPGSFQGLMHLQKLNMIQSQ----- 254
DB 242 LSFYDNLQARVPRALBOVPGKFLDKLPQVRGPGDFANMLHLKELGLNNMEELVSI 301
QY 255 -----IQVIERNAFDNQLSGLVEINLAHNHNLTLPHDLFTPLHL 293
DB 302 DKFALVNLPELTLDITNNPLSLFIHPRAFHLPQMETMLNNAALSHQOQTVESLPNL 361
QY 294 ERIHLHNPNWNCIDILWLS---HWIKDAPSNTACCARCNTPPNKGRIYIGELDQNYFT 350
DB 362 QEVGLHGNPIRCDVCIRWANATGTRVRFEIPEQSTLCA-----EPDQLQRLPVREVPFREM 417
QY 351 CYAPVIVEP---PADLNVTEGMAELKRA-STLSVSWITPNTGVTMTHGAYKVRVAVL 406
DB 418 DHCLPLISPRFSPSLQVAGESMVLHCRALAEPEPIYVWTPAGLRLTPAHAGRRCRVY 477
QY 407 SDGTLNFTNTVQDTGMVTCVNSGNTATASALTNTAATTPFSYFTVTVTETMPSQ 466
DB 478 PEGTLELRRVTAEGAGLYTCVAQNLVGADTKVSVVVGRA-----LLQPGR 523
QY 467 DEATTNNNGVTP-----VVDWETNTVTSLSLTPOSTRSTKFT 506
DB 524 DEGQGLELRQETHPYHLLSW-----VTPDNTVTSNLTWS 559

RESULT 2

CHAD_BOVIN
ID CHAD_BOVIN STANDARD; PRT; 361 AA.
AC Q27972;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Chondroadherin precursor (Cartilage leucine-rich protein) (38 kDa bone
DE protein).
GN CHAD
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
ON NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Cartilage;
RX MEDLINE=94342341; PubMed=8063792;
RA Neame P.J., Sommarin Y., Boynton R.E., Heinagaard D.;
RT "The structure of a 38-kDa leucine-rich protein (chondroadherin)
RT isolated from bovine cartilage.";
RL J. Biol. Chem. 269:21547-21554(1994).
RN [2]
RP SEQUENCE OF 25-55 AND 77-97.
RC TISSUE=Bone;
RX MEDLINE=95113864; PubMed=7814406;
RA Hu B., Coulson L., Moyer B., Price P.A.;
RT "Isolation and molecular cloning of a novel bone phosphoprotein
RT related in sequence to the cystatin family of thiol protease
RT inhibitors.";
RL J. Biol. Chem. 270:431-436(1995).
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U08018; AAA21330.1;
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00560; LRR; 10.
DR Pfam; PF01463; LRRCT; 1.
DR Pfam; PF01482; LRRNT; 1.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_TYP; 5.
KW Repeat; Signal.
FT SIGNAL 1 24 OR 23 (IN SOME ISOFORM(S)).
FT CHAIN 25 361 CHONDROADHERIN.
FT CHAIN 25 352 CHONDROADHERIN, MINOR FORM.
FT DOMAIN 79 317 10 X 24 AA LEUCINE-RICH TANDEM REPEATS.
FT REPEAT 79 102 1.
FT REPEAT 103 126 2.
FT REPEAT 127 150 3.
FT REPEAT 151 174 4.
FT REPEAT 175 198 5.
FT REPEAT 199 222 6.
FT REPEAT 223 246 7.
FT REPEAT 248 271 8.
FT REPEAT 272 293 9.
FT REPEAT 294 317 10.
FT DISULFID 306 348
FT DISULFID 308 328
FT CONFLICT 25 25 C -> Y (IN REF. 2).
FT CONFLICT 29 29 C -> W (IN REF. 2).
FT CONFLICT 31 31 C -> H (IN REF. 2).
FT CONFLICT 40 40 C -> L (IN REF. 2).
FT CONFLICT 52 52 S -> R (IN REF. 2).

SO SEQUENCE 361 AA; 40884 MW; DA79DC98AD3DD1F8 CRC64;

Query Match 12.2% Score 411.5; DB 1; Length 361;
Best Local Similarity 32.4%; Pred. No. 6.3e-21;
Matches 113; Conservative 54; Mismatches 145; Indels 37; Gaps 10;

QY 25 PLLVLLALQLLVVAGLVRA--QTCPSCVSCSNQFQSVKICVRKRLREVDPGISTNRLNL 83

Db 4 PMLLLXLSGLL--ASLLPALAACPCQCHCHSDIQLHYICDKVGLQKIPK-VSEKTKLLNL 60

QY 84 HENQIQIKVNSFK-----HURHLEI-----LQLSRNHRTIEIGA 119

Db 61 QRNFPPVLTATNSFRAMPNLVSLHLHQCOIREVAAGAFGLKQLIYLYLSHNDIRVLRAGA 120

QY 120 FNGLANLTLELFNRLTTPNGAFVYLSKLKELWLRNPNIESPSVAFNRIPSLRDL 179

Db 121 FDDTETLYLDINKTEPRGLSPLVNLFIQLANNKIRELRSGAQKDLRWLYL 180

QY 180 GELKRLSYISEGAFGLSNRYLNLAMCNLEIPN--LTPLIKLDELDSGNHLSAIRPG 237

Db 181 SE-NSLSLOPCALDDVENLAKFYLDRLNOLSSYPSSAALSCLRVVEELKLSHNPLKSIDN 239

QY 238 SFQGL-MLHQLKLMQISOQIQUIERNAFDNLQSLVEINLAHNLLPLHLETPPLHLERI 296

Db 240 AFQSGFRYLETLDNTNLEKFSDFGALGVTTLKHVLENNRLHQLPSNF--PFDSLETL 297

QY 297 HLHNPNWNCIDIWLWSWIKMADPSNTACCANCNTPNLKGRIYIGELD 345

Db 298 TLTNPWKCTQLGLRLRLLEAKTSRDPATCA---SPAKFRGQHIRDTD 343

RESULT 3

GPV_MOUSE

ID GPV_MOUSE STANDARD; PRT; 567 AA.

AC 008742;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Platelet glycoprotein V precursor (GPV) (CD42D).

GN GP5.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6; TISSUE=Liver;

RX MEDLINE=97275136; PubMed=9129030;

RA Ravanat C., Moares M., Azorsa D.O., Moog S., Schuhler S.,

RA Grunert P., Loew D., van Dorsselaer A., Cazenave J.-P., Lanza F.;

RT "Gene cloning of rat and mouse platelet glycoprotein V:

RT identification of megakaryocyte-specific promoters and demonstration

RT of functional thrombin cleavage";

RL Blood 89:3253-3262(1997).

CC -!- FUNCTION: THE GP1B-V-IX COMPLEX FUNCTIONS AS THE VON WILLEBRAND

CC FACTOR RECEPTOR AND MEDIATES VON WILLEBRAND FACTOR-DEPENDENT

CC PLATELET ADHESION TO BLOOD VESSELS. THE ADHESION OF PLATELETS TO

CC INJURED VASCULAR SURFACES IN THE ARTERIAL CIRCULATION IS A

CC CRITICAL INITIATING EVENT IN HEMOSTASIS (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- SIMILARITY: CONTAINS 14 LEUCINE-RICH REPEATS (LRR).

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; Z69595; CAA93441.1; -

DR MGD; MG1:1096363; Gp5.

DR

DR InterPro: IPR001611; LRR.
DR InterPro: IPR000483; LRR_Cterm.
DR InterPro: IPR000372; LRR_Nterm.
DR InterPro: IPR003592; LRR_out.
DR InterPro: IPR003591; LRR_Typ.
DR Pfam: PF00560; LRR; 13.
DR Pfam: PF01463; LRRCT; 1.
DR PRINTS: PR00019; LEURICHRPT.
DR SMART: SM00370; LRR; 1.
DR SMART: SM00082; LRRCT; 1.
DR SMART: SM00013; LRRNT; 1.
DR SMART: SM00369; LRR_Typ; 10.
DR Platelet; Transmembrane; Glycoprotein; Blood coagulation;
KW Repeat; Leucine-rich repeat; Cell adhesion; Signal.
FT SIGNAL 1 16
FT CHAIN 17 567
FT DOMAIN 17 522
FT TRANSMEM 523 543
FT DOMAIN 544 567
FT REPEAT 73 96
FT REPEAT 97 120
FT REPEAT 122 144
FT REPEAT 145 168
FT REPEAT 170 192
FT REPEAT 194 216
FT REPEAT 217 240
FT REPEAT 241 264
FT REPEAT 266 288
FT REPEAT 289 312
FT REPEAT 314 337
FT REPEAT 338 361
FT REPEAT 362 385
FT REPEAT 386 409
FT CARBOHYD 51 51
FT CARBOHYD 67 67
FT CARBOHYD 181 181
FT CARBOHYD 243 243
FT CARBOHYD 298 298
FT CARBOHYD 312 312
FT CARBOHYD 385 385
SQ SEQUENCE 567 AA; 63467 MW; C48643AA73967A7D CRC64;

Query Match 11.0%; Score 369.5; DB 1; Length 567;
Best Local Similarity 30.1%; Pred. No. 8e-18;
Matches 119; Conservative 63; Mismatches 166; Indels 47; Gaps 13;

QY 1 MLNKTLLHPQIMIGPRENRALEFDPDLVLLALQLLVVAGLVRAQTC---PSVCSCSNQF 57

Db 117 ILDKMVL-LEQLFLDHNALRDQNLFOOLRLQEL---GLNQNLFLPANLFSLSREL 172

QY 58 SKVICVRKNLREVDPGI---STNTRLLNHNQIQIKVNSFKHLRHLLEILQLSRNHRT 114

Db 173 KLILSRNRLTHLPKGLLGAQVKLEKLLLYSNQLTSVDSGLLSLNGALTELRNHLRS 232

QY 115 IEIGAFNGLANLTLEFNDRLTTPNGAFVYLSKLKELWLRNPNIESPSVAFNRIPSL 174

Db 233 VAPGAFDRGLNLSLTSGLNLSLPPALFLHVSSVSRLTLFENPLELPDVLFGEMAGL 292

QY 175 RRLDGLKRLSYISEGAFGLSNRYLNLAMCNLEIPNLTPLIK-----LDELDSGN 229

Db 293 RELWING-THLSTLPAAAFRLNLSGLTGLT-----RNPRLSALPRGVFOGLRLRLVGL 346

QY 230 H---LSAIRPGSFOGLMHLQKLMWIOSIOIVTERNAFNLQSLVEINLAHNLLPLHDL 286

Db 347 HTNALAEURDALLRGLHRLQVSLRHNRLRALPTLFRNLSSLSVQLHQLTLPDGV 406

QY 287 FTPLHLHLRIHLHNPNWNCIDIWLWSWIKMADPSNTACCANCNTPNLKGRIYIGELDQ 346

Db 407 FAALPQLTQVLVLLGHNPWLCDCGLWRFLQWLRLH-----HDIILGR--DEPPQ 450

QY 347 NYFTCYAPVIVEPADLNVNTEGMAELKCRASTSL 381

Db 451 ----CRGP----EPRASLFWELQGDWCPDPRSL 478

RESULT 4

GPV_RAT ID GPV_RAT STANDARD; PRT; 567 AA.

AC O08770;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Platelet glycoprotein V precursor (GPV) (CD42b).

GN GFS.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Wistar; Tissue=Liver;

RX MEDLINE=97275136; PubMed=9129030;

RA Ravanat C., Morales M., Azorsa D.O., Moog S., Schuhler S.,

RA Grunert P., Loew D., van Dorsselaer A., Cazenave J.-P., Lanza F.;

RT "Gene cloning of rat and mouse platelet glycoprotein V;

RT identification of megakaryocyte-specific promoters and demonstration

RT of functional thrombin cleavage.";

RL Blood 89:3253-3262(1997).

CC -!- FUNCTION: THE GPIB-V-IX COMPLEX FUNCTIONS AS THE VON WILLEBRAND

CC FACTOR RECEPTOR AND MEDIATES VON WILLEBRAND FACTOR-DEPENDENT

CC PLATELET ADHESION TO BLOOD VESSELS. THE ADHESION OF PLATELETS TO

CC INJURED VASCULAR SURFACES IN THE ARTERIAL CIRCULATION IS A

CC CRITICAL INITIATING EVENT IN HEMOSTASIS (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- SIMILARITY: CONTAINS 14 LEUCINE-RICH REPEATS (LRR).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; Z69594; CAA93440.1; -

DR InterPro; IPR001611; LRR.

DR InterPro; IPR000483; LRR_Cterm.

DR InterPro; IPR000372; LRR_Nterm.

DR InterPro; IPR003592; LRR_out.

DR InterPro; IPR003591; LRR_typ.

DR Pfam; PF00560; LRR; 14.

DR Pfam; PF01463; LRRCT; 1.

DR PRINTS; PR00019; LEURICHRPT.

DR SMART; SM00370; LRR; 2.

DR SMART; SM00082; LRRCT; 1.

DR SMART; SM00013; LRRNT; 1.

DR SMART; SM00369; LRR_TTP; 10.

DR Platelet; Transmembrane; Glycoprotein; Blood coagulation;

KW Repeat; Leucine-rich repeat; Cell adhesion; Signal.

FT SIGNAL 1 16

FT CHAIN 17 567

FT DOMAIN 17 522

FT TRANSMEM 523 543

FT DOMAIN 544 567

FT POTENTIAL.

FT CYTOPLASMIC (POTENTIAL).

FT LRR 1.

FT LRR 2.

FT LRR 3.

FT LRR 4.

FT LRR 5.

FT LRR 6.

FT LRR 7.

FT LRR 8.

FT LRR 9.

FT LRR 10.

FT LRR 11.

FT REPEAT 338 361 LRR 12.
FT REPEAT 362 385 LRR 13.
FT REPEAT 387 409 LRR 14.
FT CARBOHYD 51
FT CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 243 243 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 298 298 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 312 312 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 385 385 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 498 498 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 567 AA; 63344 MW; CAL0708E0D03707F CRC64;

Query Match 10.8%; Score 361.5; DB 1; Length 567;
Best Local Similarity 29.0%; Pred. No. 2.8e-17;
Matches 121; Conservative 59; Mismatches 180; Indels 57; Gaps 11;

Qy 1 MNKMTLHPQOIMIGPFRNRAFDPLLVVLLALQLLVVLVAGLVRAQTCSVCSCSQNFS-- 58
Db 117 ILDRKVL-LEQLFDHNAIRDLDQNLFOKLENLR-----DLCINQQLSFL 161
Qy 59 -----KVICV-RKNLREVPDGI---STNTRLLNLHENOIQIITKVNSEKHLRHL 103
Db 162 PANLFSSLGKLVKVDLSRNNLTHLPQGLGGAQIKLEKULLYSNRULMSDLSGLLANIGALT 221
Qy 104 ILQLSRNHIRTIEGAFNGLANLNTLEFDRNLRTTPNGAEVYLSKELMLRNNPIESI 163
Db 222 ELRLERNHLRSIAPGAFDSIGNLSTLTLSGNLLESPPALFLHYSWLTTLFENPLEEL 281
Qy 164 PSYAFNIPSRRLDGLGELKRLSYISGAFGLSMLRYNLAMCMRLRNPITP-----L 218
Db 282 PEVLFGEMAGRLRWLNG-THLRTLPAFAAFRLNSGLQTLGLTRNPL--LSALPGMFHGL 338
Qy 219 IKLDELDSGNHLSAIRPGSFQGLMHQKLMWIOSIQIIVERNADFNLQSLVEINLANN 278
Db 339 TELRLVAVHTNALBELPEDALRGGLRQVSLRHNRRLRALPTRLFRNLSLVTVLEHNQ 398
Qy 279 LTLPLPHDLFTPLHLERLHLHNPWNCNDILWLSWI-----KDMAPSNTACCARC 330
Db 399 LKTLPGDVFALPOLTRVLLGNHPWLCDCGLWPFQWLRRHLELLGRDEPP-----QC 451
Qy 331 NTPNKLGRYIGELDONFTCYAPVIVPEPPADNLNTEGMAAEKLCRASTSLTSVSWI 387
Db 452 NGPESRASLTFWELQGDQWC--PSSRGLPPDPPTENALKAPDPTQRPNSQSQSWAW 506

RESULT 5

SLIT_DROME

ID SLIT_DROME STANDARD; PRT; 1480 AA.

AC P24014;

DT 01-MAR-1992 (Rel. 21, Created)

DT 01-MAR-1992 (Rel. 21, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)

DE Slit protein precursor.

GN SLI.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=91099665; PubMed=2176636;

RA Rothberg J.M., Jacobs J.R., Goodman C.S., Artavanis-Tsakonas S.;

RT "Slit: an extracellular protein necessary for development of midline

RT glia and commissural axon pathways contains both EGF and LRR

RT domains.";

RL Genes Dev. 4:2169-2187(1990).

CC -!- FUNCTION: NECESSARY FOR DEVELOPMENT OF MIDLINE GLIA AND

CC COMMISSURAL AXON PATHWAYS. SLIT MAY INTERACT WITH EXTRACELLULAR

CC MATRIX MOLECULES.

CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A

CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.

CC -!- TISSUE SPECIFICITY: EXCRETED BY THE MIDLINE GLIA CELLS AND
CC EVENTUALLY DISTRIBUTED ALONG THE AXONS.
CC -!- SIMILARITY: CONTAINS 7 EGF-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 24 LEUCINE-RICH REPEATS (LRR).
CC -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
CC -!- SIMILARITY: CONTAINS 1 LAMININ G-LIKE DOMAIN.

CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

CC EMBL; X53959; CAA37910.1; -

DR PIR; A36665; A36665.

DR HSP; P00743; IICF.

DR FlyBase; FBgn003425; sli.

DR InterPro; IPR000152; Asx_hydroxyl.

DR InterPro; IPR000359; Cys_knot.

DR InterPro; IPR000561; EGF-like.

DR InterPro; IPR000742; EGF-2.

DR InterPro; IPR001881; EGF_Ca.

DR InterPro; IPR001611; LRR.

DR InterPro; IPR000483; LRR_Cterm.

DR InterPro; IPR000372; LRR_Nterm.

DR InterPro; IPR003592; LRR_Out.

DR InterPro; IPR003591; LRR_typ.

DR InterPro; IPR001791; Laminin_G.

DR Pfam; PF00007; Cys_knot; 1.

DR Pfam; PF00008; EGF; 7.

DR Pfam; PF00034; laminin_G; 1.

DR Pfam; PF00560; LRR; 17.

DR Pfam; PF01463; LRRCT; 4.

DR Pfam; PF01462; LRRNT; 4.

DR SMART; SM00041; Ct; 1.

DR SMART; SM000179; EGF_Ca; 2.

DR SMART; SM000001; EGF_like; 5.

DR SMART; SM00370; LRR; 4.

DR SMART; SM00082; LRRCT; 4.

DR SMART; SM00013; LRRNT; 4.

DR SMART; SM00369; LRR_Typ; 9.

DR SMART; SM00282; Lamg; 1.

DR PROSITE; PS00010; ASX_HYDROXYL; 3.

DR PROSITE; PS01185; CTCK_1; 1.

DR PROSITE; PS01225; CTCK_2; 1.

DR PROSITE; PS00022; EGF_1; 7.

DR PROSITE; PS01186; EGF_2; 5.

DR PROSITE; PS01187; EGF_Ca; 2.

DR PROSITE; PS00025; LAM_G_DOMAIN; 1.

DR Neurogenesis; Glycoprotein; Signal; Alternative splicing;

DR EGF-like domain; Repeat; Leucine-rich repeat.

DR SIGNAL 1 36

FT CHAIN 37 1480 SLIT PROTEIN.

FT REPEAT 99 122 LRR 1.

FT REPEAT 123 146 LRR 2.

FT REPEAT 148 170 LRR 3.

FT REPEAT 171 194 LRR 4.

FT REPEAT 195 218 LRR 5.

FT REPEAT 220 246 LRR 6.

FT REPEAT 321 344 LRR 7.

FT REPEAT 345 368 LRR 8.

FT REPEAT 369 392 LRR 9.

FT REPEAT 394 416 LRR 10.

FT REPEAT 417 440 LRR 11.

FT REPEAT 522 545 LRR 12.

FT REPEAT 546 569 LRR 13.

FT REPEAT 570 593 LRR 14.

FT REPEAT 595 617 LRR 15.

FT REPEAT 618 641 LRR 16.

FT REPEAT 643 666 LRR 17.

FT REPEAT 678 701 LRR 18.

FT REPEAT 720 743
FT REPEAT 745 764 LRR 19.
FT REPEAT 765 788 LRR 20.
FT REPEAT 790 812 LRR 21.
FT REPEAT 813 836 LRR 22.
FT REPEAT 838 861 LRR 23.
FT REPEAT 907 944 LRR 24.
FT DOMAIN 946 983 EGF-LIKE 1.
FT DOMAIN 985 1022 EGF-LIKE 2.
FT DOMAIN 1024 1062 EGF-LIKE 3.
FT DOMAIN 1064 1100 EGF-LIKE 4.
FT DOMAIN 1111 1149 EGF-LIKE 5.
FT DOMAIN 1152 1325 LAMININ G-LIKE.
FT DOMAIN 1353 1392 EGF-LIKE 6.
FT DOMAIN 1409 1480 EGF-LIKE 7.
FT CARBOHYD 111 111 CTCK.
FT CARBOHYD 207 207 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 357 357 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 435 435 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 783 783 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 788 788 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 958 958 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 998 998 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1060 1060 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1159 1159 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1175 1175 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1243 1243 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1292 1292 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 911 922 BY SIMILARITY.
FT DISULFID 916 932 BY SIMILARITY.
FT DISULFID 934 943 BY SIMILARITY.
FT DISULFID 950 961 BY SIMILARITY.
FT DISULFID 955 971 BY SIMILARITY.
FT DISULFID 973 982 BY SIMILARITY.
FT DISULFID 989 1001 BY SIMILARITY.
FT DISULFID 995 1010 BY SIMILARITY.
FT DISULFID 1012 1021 BY SIMILARITY.
FT DISULFID 1028 1041 BY SIMILARITY.
FT DISULFID 1035 1050 BY SIMILARITY.
FT DISULFID 1052 1061 BY SIMILARITY.
FT DISULFID 1068 1079 BY SIMILARITY.
FT DISULFID 1073 1088 BY SIMILARITY.
FT DISULFID 1090 1099 BY SIMILARITY.
FT DISULFID 1115 1125 BY SIMILARITY.
FT DISULFID 1120 1137 BY SIMILARITY.
FT DISULFID 1139 1148 BY SIMILARITY.
FT DISULFID 1357 1368 BY SIMILARITY.
FT DISULFID 1362 1380 BY SIMILARITY.
FT DISULFID 1382 1391 BY SIMILARITY.
FT DISULFID 1409 1443 BY SIMILARITY.
FT DISULFID 1423 1457 BY SIMILARITY.
FT DISULFID 1434 1473 BY SIMILARITY.
FT DISULFID 1438 1475 BY SIMILARITY.
FT DISULFID 1442 1479 BY SIMILARITY.
FT VARSPLIC 1394 1404 MISSING (IN SHORT ISOFORM).
SQ SEQUENCE 1480 AA; 165752 MW; F9D5925FC170B1C3 CRC64;

Query Match 10.4% Score 350; DB 1; Length 1480;

Best Local Similarity 23.9% Pred. No. 5.8e-16;

Matches 109; Conservative 65; Mismatches 152; Indels 130; Gaps 9;

Qy 46 TCPSCVCSNFGSKVICVRKNLREVPDGIISTNRLNLHENQIQIKVNSFKHLRLEL 105

Db 294 SCPPRCADGI--VDCREKSLTSVPVTLPPDITDVRLEQNFITELPKFSFSFRRLRI 351

Qy 106 QLSNRHRTIEIGAPFNLANLTLELFDNRLTTPNGAFVYLSKULKELWLRNNTIESIPS 165

Db 352 DLSNNNISRIADALSGLKQLTTLVLYGNKIKDLPSPGVFKGLSLRLLLANEISCIRK 411

Qy 166 YAFNRIPSLRRDLGELKRLSYISEGAPEGLSNRLYLNAM-----CNLR----- 210

Db 412 DAFRLHSLSLSLYD--NNIOSLANGTFDAMKSMKTVHLAKNPFCDNLRWLADYLNHN 470

Qy 211 -----EIPNLPL-----TKLD-----ELDSGNHLSAIRPGS 238
Db 471 PIETSGARCESPKRMHRRRIEBSLREKFKCSWGLRMKLSGECRMDSDCPAMCHCEGTVV 530
Qy 211 -----EIPNLPL-----TKLD-----ELDSGNHLSAIRPGS 238
Db 531 DCTGRLKEIPDIPLHTTELLNDNELGRISDGLFGLRPLHLVKLELKRNLQTGIEPNA 590
Qy 239 FQGLMHLOKLMQISOQVIERNAFDNLQSLVEINLAHNNLTLLPHDLFTPLHLERHL 298
Db 591 FEGASHIQELQGENKIKETSINKMFLGLHQLKTLNLYDNQISCVMPGSEFHLNLSLNL 650
Qy 299 HHNPWNCNDLILSWIKDMAPSNATCCARCNTPPNLKGRYIGELDONVFTCYAPVIVE 358
Db 651 ASNFENCNCHLAWFAECVRKKSLLNGA--ARCGAPSKVRDVOIKDLPHSEFKCSS----- 703
Qy 359 PPADLVTEGMAELKCRASLTSLTSVSWITPNTGTVM 394
Db 704 -----ENSEGCGDGYCPPSCTCT-----GTWV 726

RESULT 6
ALS_PAPHA ID ALS_PAPHA STANDARD; PRT; 605 AA.
AC 002833;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Insulin-like growth factor binding protein complex acid labile chain precursor (ALS).
GN IGFALS OR ALS.
OS Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecoidea; Papio.
OC NCBI_TaxID=9557;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Liver;
RX MEDLINE=97040714; PubMed=8886027;
RA Delhanty P., Baxter R.C.;
RT "The cloning and expression of the baboon acid-labile subunit of the insulin-like growth factor binding protein complex.";
RL Biochem. Biophys. Res. Commun. 227:897-902(1996).
CC -!- FUNCTION: INVOLVED IN PROTEIN-PROTEIN INTERACTIONS THAT RESULT IN PROTEIN COMPLEXES, RECEPTOR-LIGAND BINDING OR CELL ADHESION.
CC -!- SUBUNIT: FORMS A TERNARY COMPLEX OF ABOUT 140 TO 150 kDa WITH IGF-I OR IGF-II AND IGF-BP-3 (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- SIMILARITY: CONTAINS 20 LEUCINE-RICH REPEATS (LRR).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S83462; -. NOT_ANNOTATED_CDS.
CC HSSP; P23945; 1XUN.
CC InterPro; IPR001611; LRR.
CC InterPro; IPR000483; LRR_term.
CC InterPro; IPR000372; LRR_Nterm.
CC InterPro; IPR003592; LRR_out.
CC InterPro; IPR003591; LRR_typ.
CC Pfam; PF00560; LRR; 19.
CC Pfam; PF01463; LRRCT; 1.
CC PRINTS; PR00019; LEURICHRPT.
CC SMART; SM00370; LRR; 2.
CC SMART; SM00082; LRRCT; 1.

DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_Typ; 11.
KW Glycoprotein; Leucine-rich repeat; Repeat; Signal.
FT SIGNAL 1 27 BY SIMILARITY.
FT CHAIN 28 605 INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN COMPLEX ACID LABILE CHAIN.
FT REPEAT 53 73 LRR 1.
FT REPEAT 74 96 LRR 2.
FT REPEAT 98 120 LRR 3.
FT REPEAT 121 144 LRR 4.
FT REPEAT 146 168 LRR 5.
FT REPEAT 169 192 LRR 6.
FT REPEAT 194 216 LRR 7.
FT REPEAT 217 240 LRR 8.
FT REPEAT 242 264 LRR 9.
FT REPEAT 265 288 LRR 10.
FT REPEAT 289 312 LRR 11.
FT REPEAT 313 336 LRR 12.
FT REPEAT 338 360 LRR 13.
FT REPEAT 361 384 LRR 14.
FT REPEAT 386 408 LRR 15.
FT REPEAT 409 432 LRR 16.
FT REPEAT 433 456 LRR 17.
FT REPEAT 458 480 LRR 18.
FT REPEAT 482 504 LRR 19.
FT REPEAT 505 530 LRR 20.
FT CARBOHYD 64 64 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 85 85 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 96 96 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 368 368 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 515 515 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 580 580 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 605 AA; 66110 MW; 9D71994625F23652 CRC64;

Query Match 10.0%; Score 335; DB 1; Length 605;
Best Local Similarity 20.8%; Pred. No. 1.9e-15;
Matches 125; Conservative 64; Mismatches 125; Indels 286; Gaps 13;

Qy 31 LALQLLVVA-----GLVRAQTCPSCVSCS-----NOFSCVICVRKRLR 68
Db 8 LALALLLSWALGRSLRLEGAEPGTPGEAGPACATCATCSDYDEVNELS-VFCSRNLT 66
Qy 69 EVDGISTNTRL-----
Db 67 RLPDGIPTGTOALWLDSDNNLSIPPAFRNLSSLAFLNLOGGQLGSLPEQALLGLENLCH 126
Qy 81 LNLHENQIITKVSFKHLRHLLEILQLSRNHRTIEICAFNGLANLNTLELFDNRLTTP 140
Db 127 LHLERNQLRSVAVGTFTAYTPALALLGLSNNRLSRLEDGLFEGGLNLDNLGNWNLAVLP 186
Qy 141 NGAF-----VY-----LSKLKELWLRNNPIESIPSYAFNRIPSLRR 176
Db 187 DAARFGLGLRELVLGNRLAYLQALPFLSGLAELRELDLSRNALRAIKANVFAQLPRLOK 246
Qy 177 L-----
Db 247 LYLDRLNIAAVAPGAFGLKALRWLDLSHNRVAGLLETFPCLLGLRLVRLSLHNAIASLR 306
Qy 178 -----DLGELK-----RLSYISGEGAFGLSLRLY 201
Db 307 PRTFDLHFLLEELQIGHNRIRIQLARSPRSGEGVQLLEVLTDHNLQGVKVFAGLUTNVAV 366
Qy 202 LNLAMCNLRIPN--LTPILKLDLDSGNHLSAIRPGSFGQGLMHLOKLMWISQIOVIE 259
Db 367 MNLSCNLRNLPEQVFRGLGKLSLHLSGLGRIRPHTFAGLSGLRLFLKDKNLGVIE 426
Qy 260 RNAPDNQSLVEINLAHNNLTLLPHDLFTPL-----
Db 427 EQSLWGLAELELDLTSNQLTHLPQLFQGLKLEKLELHLLSHNRLAEALPADALGPLQRAFW 486
Qy 291 ---HH-----LERIHLHNPNCNCIDILW 311

Db 487 LDVSHNRLKALPSGLSLASGLRLRYLNLNRNLSLRTFTPPQPLERLMLGNPWCSCPLKA 546;
 QY 312 LSWIKMDAPNTACAR-----CNTPPNLKGRYIGELDONYE 349;
 Db 547 L-----RDFALONSAVPRFVQAIPEGDCOPVYTYNNITCASPPVAGLDLRLDGEAHF 602

RESULT 7
 ID GPV_HUMAN STANDARD; PRT; 560 AA.
 AC PA0197;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Platelet glycoprotein V precursor (GPV) (CD42D).
 GN GP5.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=93391348; PubMed=7690959;
 RA Hickey M.J., Hagen F.S., Yagi M., Roth G.J.;
 RT "Human platelet glycoprotein V: characterization of the polypeptide
 and the related ib-v-ix receptor system of adhesive, leucine-rich
 glycoproteins.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:8327-8331(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=platelet;
 RX MEDLINE=94012616; PubMed=8407908;
 RA Lanza F., Morales M., de la Salle C., Cazenave J.-P., Clemetson K.J.,
 RA Shimomura T., Phillips D.R.;
 RT "Cloning and characterization of the gene encoding the human platelet
 glycoprotein V. A member of the leucine-rich glycoprotein family
 cleaved during thrombin-induced platelet activation.";
 RL J. Biol. Chem. 268:20801-20807(1993).
 RN [3]
 RP PARTIAL SEQUENCE.
 RC TISSUE=platelet;
 RX MEDLINE=90275263; PubMed=2350580;
 RA Shimomura T., Fujimura K., Maehama S., Takemoto M., Oda K.,
 RA Fujimoto T., Oyama R., Suzuki M., Ichiara-Tanaka K., Titani K.,
 RA Kuramoto A.;
 RT "Rapid purification and characterization of human platelet
 glycoprotein V: the amino acid sequence contains leucine-rich
 repetitive modules as in glycoprotein Ib.";
 RL Blood 75:2349-2356(1990).
 RN [4]
 RP PARTIAL SEQUENCE.
 RC TISSUE=platelet;
 RX MEDLINE=90321220; PubMed=2372284;
 RA Roth G.J., Church T.A., McMullen B.A., Williams S.A.;
 RT "Human platelet glycoprotein V: a surface leucine-rich glycoprotein
 related to adhesin.";
 RL Biochem. Biophys. Res. Commun. 170:153-161(1990).
 CC -1- FUNCTION: THE GPIB-V-IX COMPLEX FUNCTIONS AS THE VON WILLEBRAND
 FACTOR RECEPTOR AND MEDIATES VON WILLEBRAND FACTOR-DEPENDENT
 PLATELET ADHESION TO BLOOD VESSELS. THE ADHESION OF PLATELETS TO
 INJURED VASCULAR SURFACES IN THE ARTERIAL CIRCULATION IS A
 CRITICAL INITIATING EVENT IN HEMOSTASIS.
 CC -1- SUBCELLULAR LOCATION: TYPE I membrane protein.
 CC -1- TISSUE SPECIFICITY: PLATELETS AND MEGAKARYOCYTES.
 CC -1- PTM: THE N-TERMINUS IS BLOCKED.
 CC -1- SIMILARITY: CONTAINS 14 LEUCINE-RICH REPEATS (LRR).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L11238; AAA03069.1; -;
 DR EMBL: Z23091; CAA80637.1; -;
 DR HSSP: P16473; IUXM.
 DR MIM: 173511; -;
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000483; LRR_Cterm.
 DR InterPro: IPR000372; LRR_Nterm.
 DR InterPro: IPR003592; LRR_Out.
 DR InterPro: IPR003591; LRR_typ.
 DR Pfam: PF00560; LRR; 14.
 DR Pfam: PF01463; LRRCT; 1.
 DR PRINTS: PR00019; LEURICHRPT.
 DR SMART: SM00370; LRR; 1.
 DR SMART: SM00082; LRRCT; 1.
 DR SMART: SM00013; LRRNT; 1.
 DR SMART: SM00369; LRR_typ; 9.
 DR Platelet; Transmembrane; Glycoprotein; Blood coagulation;
 KW Repeat; Leucine-rich repeat; Cell adhesion; Signal.
 FT SIGNAL 1 16
 FT CHAIN 17 560 PLATELET GLYCOPROTEIN V.
 FT DOMAIN 17 523 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 524 544 POTENTIAL.
 FT DOMAIN 545 560 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 73 96 LRR 1.
 FT REPEAT 97 120 LRR 2.
 FT REPEAT 122 144 LRR 3.
 FT REPEAT 145 168 LRR 4.
 FT REPEAT 169 192 LRR 5.
 FT REPEAT 194 216 LRR 6.
 FT REPEAT 217 240 LRR 7.
 FT REPEAT 241 264 LRR 8.
 FT REPEAT 266 288 LRR 9.
 FT REPEAT 289 312 LRR 10.
 FT REPEAT 314 337 LRR 11.
 FT REPEAT 338 361 LRR 12.
 FT REPEAT 362 385 LRR 13.
 FT REPEAT 386 409 LRR 14.
 FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 181 181 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 243 243 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 267 267 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 298 298 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 312 312 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 385 385 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 499 499 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 73 74 MT -> TK (IN REF. 2).
 FT CONFLICT 109 109 K -> T (IN REF. 2).
 FT CONFLICT 130 130 D -> W (IN REF. 3).
 FT CONFLICT 136 138 GID -> PGG (IN REF. 3).
 FT CONFLICT 209 209 L -> I (IN REF. 2).
 FT CONFLICT 267 267 N -> H (IN REF. 3).
 FT CONFLICT 327 327 L -> I (IN REF. 2).
 FT CONFLICT 478 478 P -> G (IN REF. 2).
 FT CONFLICT 509 509 P -> D (IN REF. 2).
 SQ SEQUENCE 560 AA; 60959 MW; BICDB04AF8AF7115 CRC64;

Query Match 9.8%; Score 329.5; DB 1; Length 560;
 Best Local Similarity 29.7%; Pred. No. 4e-15;
 Matches 83; Conservative 43; Mismatches 104; Indels 49; Gaps 3;
 QY 66 NLREVDPG-----1STNTRLLN-----LHENQIQIKVNSFKH 98
 Db 181 NLTHLPKGLGAQAKLERLLHSHNRLVSLDGLNLSGALTELOFHRNHSIAPGAEDR 240
 QY 99 LRHEILQLSRNHRITTEIGAFNGLANLTLEFDNRUTTIIPNGAFVYLSKLKELWLRNN 158
 Db 241 LPNLSSTLSNRHLAFPSALFLSHNLTLTLPENPAELPGVLFGEHGGGLQELWLRNT 300
 QY 159 PIESIPSYAFNRIPSLRRRLDGLKRLSYISEGAFEGLSNRLYLNLMCMNLRIPNLTP 218

```
Db 301 QURTPAAAFNRNLSRLRYGLVTSPLRSALPQGAFOGLGELQV-----343
Qy 219 IKLDELDSGNHLSAIRGSGFQGLMHLQKLMWISQIOVIERNAPDNLOSIVEINLAHNN 278
Db 344 -----LALHNSGLTLPGLLGLGKLRQVSLRRNRRLRALPRALFRALSSLESVOLHDNQ 398
Qy 279 LTLPHDLFTPLHLERHILHHPNPNWNCNDILWLSWLIK 317
Db 399 LETLPDGVFGALPRLTVEVLGHNSWRCDGLGPFGLGNLR 437

RESULT 8
ALS_HUMAN
AC 35858; STANDARD; PRT; 605 AA.
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Insulin-like growth factor binding protein complex acid labile chain
DE precursor (ALS).
GN IGFALS OR ALS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN [2]
RN [3]
RC SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RP TISSUE=Liver;
RX MEDLINE=92357025; PubMed=1379671;
RA Leong S.R., Baxter R.C., Camerato T., Dai J., Wood W.I.;
RT "Structure and functional expression of the acid-labile subunit of
RT the insulin-like growth factor-binding protein complex.";
RN Mol. Endocrinol. 6:870-876(1992).
RP [2]
RP SEQUENCE FROM N.A.
RA Frankland J.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE OF 28-35.
RX MEDLINE=89308584; PubMed=2473065;
RA Baxter R.C., Martin J.L., Beniac V.A.;
RT "High molecular weight insulin-like growth factor binding protein
RT complex. Purification and properties of the acid-labile subunit from
RT human serum.";
RN J. Biol. Chem. 264:11843-11848(1989).
CC -1- FUNCTION: INVOLVED IN PROTEIN-PROTEIN INTERACTIONS THAT RESULT
CC IN PROTEIN COMPLEXES, RECEPTOR-LIGAND BINDING OR CELL ADHESION.
CC -1- SUBUNIT: FORMS A TERNARY COMPLEX OF ABOUT 140 TO 150 kDa WITH
CC IGF-I OR IGF-II AND IGFBP-3.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: PLASMA.
CC -1- SIMILARITY: CONTAINS 20 LEUCINE-RICH REPEATS (LRR).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/annouce/
CC or send an email to license@isb-sib.ch).
```

```
DR Pfam; PF01463; LRRCT; 1.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00370; LRR; 2.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_Typ; 11.
KW Glycoprotein; Leucine-rich repeat; Repeat; Signal.
FT SIGNAL 1 27
FT CHAIN 28 605
FT REPEAT 53 73
FT REPEAT 74 96
FT REPEAT 98 120
FT REPEAT 121 144
FT REPEAT 145 168
FT REPEAT 170 192
FT REPEAT 193 216
FT REPEAT 217 240
FT REPEAT 242 264
FT REPEAT 265 288
FT REPEAT 289 312
FT REPEAT 313 336
FT REPEAT 337 360
FT REPEAT 361 384
FT REPEAT 386 408
FT REPEAT 409 432
FT REPEAT 433 456
FT REPEAT 458 480
FT REPEAT 482 504
FT REPEAT 505 530
FT CARBOHYD 64 64 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 85 85 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 96 96 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 368 368 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 515 515 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 580 580 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 605 AA; 66034 MW; F6562A23CBE918F6 CRC64;

Query Match 9.58; Score 320; DB 1; Length 605;
Best Local Similarity 29.8%; Pred. No. 2e-14;
Matches 95; Conservative 50; Mismatches 124; Indels 50; Gaps 6;

Qy 31 LALQLLVVA-----GLVRAQTCPSCVSCS-----NOFSKVICVRKNLR 68
Db 8 LALALLLSWALGPRLEGADPGTGEAGPACPAACVCSYDDADELS-VFCSSRLMT 66
Qy 69 EYDGIPTNTLLNLHENQIOIKVNSFKHLRHLLEI-----104
Db 67 RLDPGVPGGTQALWLDGNLSSVPPAFAFQNLSSGLFLNLQGGQLGSLPQALLGLENLCH 126
Qy 105 LQLSNRHIRTIEICAFNGLANLNTLELFDNRLTTPNGAFVYLSKKELWLRNPIESIP 164
Db 127 LHLENRQLRSALGTFAHTPALASGLSNRSLRLEDGLFEGLSLWDLNLGWSLAVLP 186
Qy 165 SYAFNRPLSLRDLGELKRLSYSEGAFCGLSNRLYLNLAMCNLRTPN--LPLIKLD 222
Db 187 DAAPFGLSRELVLG--NRUAYLQPAFLFSGLAELRELDLSRNALRAIKANVFOLPRQ 245
Qy 223 ELDSGNHLSAIRGSGFQGLMHLQKLMWISQIOVIERNAPDNLOSIVEINLAHNNLTLL 282
Db 246 KLYLDRNLIAAVAPGAFGLGKLRQVSLRRNRRLRALPRALFRALSSLESVOLHDNQ 305
Qy 283 PHDLFTPLHLERHILHHPN 301
Db 306 RPRTFKDLHFLLEELQGLHN 324

RESULT 9
ALS_RAT
ID ALS_RAT
AC P35859; STANDARD; PRT; 603 AA.
```


01-JUN-1994 (Rel. 29, Created)
 01-JUN-1994 (Rel. 29, Last sequence update)
 16-OCT-2001 (Rel. 40, Last annotation update)
 Insulin-like growth factor binding protein complex acid labile chain precursor (ALS).
 IGFALS OR ALS.
 Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxID=10116;
 [1]
 SEQUENCE FROM N.A.
 RP TISSUE=Liver;
 RC MEDLINE=93038676; PubMed=1384485;
 RA Dai J., Baxter R.C.;
 RA "Molecular cloning of the acid-labile subunit of the rat insulin-like growth factor binding protein complex."
 RT Biochem. Biophys. Res. Commun. 188:304-309(1992).
 RL [2]
 RP SEQUENCE OF 24-44, AND CHARACTERIZATION.
 RC STRAIN=Wistar; TISSUE=Serum;
 RX MEDLINE=94130835; PubMed=7507839;
 RA Baxter R.C., Dai J.;
 RA "Purification and characterization of the acid-labile subunit of rat serum insulin-like growth factor binding protein complex."
 RT Endocrinology 134:848-852(1994).
 RL CIRCULATING IGFs TO THE TISSUES.
 CC -1- FUNCTION: MAY HAVE AN IMPORTANT ROLE IN REGULATING THE ACCESS OF
 CC CIRCULATING IGFs TO THE TISSUES.
 CC -1- SUBUNIT: FORMS A TERNARY COMPLEX OF ABOUT 140 TO 150 kDa WITH
 CC IGF-I OR IGF-II AND IGFBP-3.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: BRAIN, KIDNEY, LUNG, HEART, SPLEEN, MUSCLE
 CC AND LIVER.
 CC -1- SIMILARITY: CONTAINS 21 LEUCINE-RICH REPEATS (LRR).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; S46785; AAB23770.2; -
 CC PIR; JC1282; JC1282.
 CC HSSP; P23945; 1XUN.
 CC InterPro; IPR001611; LRR.
 CC InterPro; IPR000483; LRR_Cterm.
 CC InterPro; IPR000372; LRR_Nterm.
 CC InterPro; IPR003592; LRR_Out.
 CC InterPro; IPR003591; LRR_Typ.
 CC Pfam; PF00560; LRR; 19.
 CC Pfam; PF01463; LRRCT; 1.
 CC Pfam; PF01462; LRRNT; 1.
 CC PRINTS; PR00019; LEURCHRT.
 CC SMART; SM00370; LRR; 5.
 CC SMART; SM00082; LRRCT; 1.
 CC SMART; SM00013; LRRNT; 1.
 CC SMART; SM00369; LRR_Typ; 9.
 KW Glycoprotein; Leucine-rich repeat; Repeat; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 603
 FT REPEAT 52 73
 FT REPEAT 74 96
 FT REPEAT 98 120
 FT REPEAT 121 144
 FT REPEAT 146 168
 FT REPEAT 169 192
 FT REPEAT 194 216
 FT REPEAT 217 240
 FT REPEAT 242 264
 FT REPEAT 266 288

FT REPEAT 289 312 LRR 11.
 FT REPEAT 313 336 LRR 12.
 FT REPEAT 337 360 LRR 13.
 FT REPEAT 361 384 LRR 14.
 FT REPEAT 386 408 LRR 15.
 FT REPEAT 409 432 LRR 16.
 FT REPEAT 433 456 LRR 17.
 FT REPEAT 458 478 LRR 18.
 FT REPEAT 479 504 LRR 19.
 FT REPEAT 506 529 LRR 20.
 FT REPEAT 543 566 LRR 21.
 FT CARBOHYD 64 64 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 85 85 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 96 96 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 368 368 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 515 515 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 578 578 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 586 586 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 603 AA; 66811 MW; DCD7637D94A5037C CRC64;
 Query Match 9.3%; Score 312; DB 1; Length 603;
 Best Local Similarity 30.0%; Pred. No. 6.8e-14;
 Matches 95, Conservative 46; Mismatches 136; Indels 40; Gaps 5;
 Qy 25 PLLVLLALQLLV-----VAGLVRAQTCPSCVSCSN-----QFSKVICVRKNLRE 69
 Db 8 PALVLLAFVWALGCHLQGTDPGASADAEGPQCPVACTCSHDDYTDLSVFCSSKNLTH 67
 Qy 70 VPDGISTNRLNLHENOIQIKVNSFKHLRHLTQLSRNHIRTEIGAFNGLANLNTL 129
 Db 68 LPDDIPVSTRALWLDGNNLSSIPSAFQNLSSLDLFLNLQGSWLSRLEPQALLGLQNLVYL 127
 Qy 130 ELFDNRLTTPNGAFVYLKELWLRNNPIESIPSYAFNIPSLRRLLDLG----- 180
 Db 128 HLERNLRLNVLGTHTPPSLASLSLSSNLGLRLEEGFLQGLSHLDNLGNLSLVLPD 187
 Qy 181 -----EL-----KRLSYISEGAFEGLSNRLYNLNLAMCNLEIPN--LPPLIKLDEL 224
 Db 188 TVFGGLGNLHELVLVAGNLTVLQPALFCGLGELRELDLSRNALRSVKANVFVHLPRQLKL 247
 Qy 225 DLSGNHLSAIRPGSQGLMHLOKLMWIOSQTOVIERNADFNLQSLVEINLAINNTLLPH 284
 Db 248 YLDNRLITAVAPGAFGLMKALRWLDLSHNRVAGLMEDTFPGLGLHVLRLAHNAIASLRP 307
 Qy 285 DLFTPLHHLERHHLHN 301
 Db 308 RTFRDLHFLLEQLQGHN 324
 RESULT 10
 ALS_MOUSE
 ID ALS_MOUSE STANDARD; PRT; 603 AA.
 AC P70389;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Insulin-like growth factor binding protein complex acid labile chain precursor (ALS).
 DE IGFALS OR ALS OR ALBS.
 GN Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SV;
 RX MEDLINE=96413591; PubMed=8816745;
 RA Boisclair Y.R., Seto D., Hsieh S., Hurst K.R., Ooi G.T.;
 RA "Organization and chromosomal localization of the gene encoding the mouse acid labile subunit of the insulin-like growth factor binding protein complex."
 RT Proc. Natl. Acad. Sci. U.S.A. 93:10028-10033(1996).
 RL

- I- FUNCTION: MAY HAVE AN IMPORTANT ROLE IN REGULATING THE ACCESS OF CIRCULATING IGFs TO THE TISSUES.
- I- SUBUNIT: FORMS A TERNARY COMPLEX OF ABOUT 140 TO 150 kDa WITH IGF-I OR IGF-II AND IGFBP-3 (BY SIMILARITY).
- I- SUBCELLULAR LOCATION: Extracellular.
- I- SIMILARITY: CONTAINS 21 LEUCINE-RICH REPEATS (LRR).

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; U66900; AAB17270.1; -
MGD; MGI:107973; Igfals.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003592; LRR_out.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00560; LRR; 19.
DR Pfam; PF01463; LRRCT; 1.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00370; LRR; 4.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_TYP; 9.
KW Glycoprotein; Leucine-rich repeat; Repeat; Signal.
FT SIGNAL 1 23
FT CHAIN 24 603
FT REPEAT 52 73
FT REPEAT 74 96
FT REPEAT 98 120
FT REPEAT 121 144
FT REPEAT 146 168
FT REPEAT 169 192
FT REPEAT 194 216
FT REPEAT 217 240
FT REPEAT 242 264
FT REPEAT 265 288
FT REPEAT 289 312
FT REPEAT 313 336
FT REPEAT 338 360
FT REPEAT 361 384
FT REPEAT 386 408
FT REPEAT 409 432
FT REPEAT 433 456
FT REPEAT 458 480
FT REPEAT 482 504
FT REPEAT 505 529
FT REPEAT 543 566
FT REPEAT 64 64
FT CARBOHYD N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 603 AA; 66959 MW; 4A6223AEC7026BCC CRC64;

Query Match 9.3%; Score 311; DB 1; Length 603;
Best Local Similarity 30.5%; Pred. No. 8e-14;
Matches 100; Conservative 44; Mismatches 122; Indels 62; Gaps 8;
OY 25 PLLVLLALQLLV-----VAGLVRAQTCPSCVSCS-----NOFSKVICVRKNLR 68
DB 8 PALVLLAFWALGPCYLGQIDPGASDAEGPQCPVTCTCSYDDYDDELVS-VFCSSRLMT 66
| | | | | : : : : : | | | | : : : : : | | | : : : : : | | : : : : : |

OY 69 EVPDGIQSTNRLNHNQIIVKNSFKHLRHLLEIQLSRNHRHRTIEIGAFNGLANLNT 128
: : : : : | | | | : : : : : | | | | : : : : : | | | | : : : : : | | | | : : : : : |
DB 67 QLPDIPVSTRALWLDGNNLSIPSAFQNLSSDLFLNQLGSWLSRSLERQALLGLQNLVH 126
: : : : : | | | | : : : : : | | | | : : : : : | | | | : : : : : | | | | : : : : : |
OY 129 LELFQNRUTTPNGAFVYLSKLKELWLRNNPIESTPSYAFNRIPSLRRDLGELKRLSYI 188
: : : : : | | | | : : : : : | | | | : : : : : | | | | : : : : : | | | | : : : : : |
DB 127 LHLENNLLRSAAAGLFRUTPSLASLGNLGNLGRLEEGFLRGLSHLWDLNLG-WNSLVVL 185
: : : : : | | | | : : : : : | | | | : : : : : | | | | : : : : : | | | | : : : : : |
OY 189 SEGAFEGISNRYLNLA-----MCNLRLEIPNLTPPLKLDLDELDSGNHLSAIRPG 237
: : : : : | | | | : : : : : | | | | : : : : : | | | | : : : : : | | | | : : : : : |
DB 186 PTFVQGGNGNHELVLGKNTLYLQPALLCGLGE-----LRELDLSRNALRSVKAN 236
: : : : : | | | | : : : : : | | | | : : : : : | | | | : : : : : | | | | : : : : : |
OY 238 SFOGLMHQLKLMQISQIQVIERNAF-----DNIQSLVE-----IN 273
: : : : : | | | | : : : : : | | | | : : : : : | | | | : : : : : | | | | : : : : : |
DB 237 VFHLPRLQKLYLDRNLITAVAPRAFLGMKALRWLDLSHNRVAGLLEDFPGLGLGLHLVR 296
: : : : : | | | | : : : : : | | | | : : : : : | | | | : : : : : | | | | : : : : : |
OY 274 LAHNNTLLPHDLFTPLHHLERIHLLHN 301
: : : : : | | | | : : : : : | | | | : : : : : | | | | : : : : : | | | | : : : : : |
DB 297 LAHNAITSLRPRTEKDLHFLLEQLQGHN 324
: : : : : | | | | : : : : : | | | | : : : : : | | | | : : : : : | | | | : : : : : |
RESULT 11
PGS2_HUMAN
ID PGS2_HUMAN STANDARD; PRT; 359 AA.
AC P07585; Q9Y5N9; Q9Y5N8; Q9P020; Q9P021;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Bone proteoglycan II precursor (PG-S2) (decorin) (PG40).
GN DCN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87017013; PubMed=3484330;
RT "Primary structure of an extracellular matrix proteoglycan core protein deduced from cloned cDNA."
RL Proc. Natl. Acad. Sci. U.S.A. 83:7683-7687(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=93162643; PubMed=8432527;
RA Vetter U., Vogel W., Just W., Young M.F., Fisher L.W.;
RT "Human decorin gene: intron-exon junctions and chromosomal localization."
RL Genomics 15:161-168(1993).
RN [3]
RP SEQUENCE OF 1-70 FROM N.A.
RX MEDLINE=93162642; PubMed=8432526;
RA Danielson K.G., Fazzio A., Cohen I.R., Cannizzaro L., Iozzo R.V.;
RT "The human decorin gene: intron-exon organization, discovery of two alternatively spliced exons in the 5' untranslated region, and mapping of the gene to chromosome 12q23."
RL Genomics 15:146-160(1993).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORMS A; B; C; D AND E).
RA Cs-Szabo G., Glant T.T.;
RT "Alternative splicing of human decorin."
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 31-50.
RX MEDLINE=90073579; PubMed=2590169;
RA Roughley P.J., White R.J.;
RT "Dermatan sulphate proteoglycans of human articular cartilage. The properties of dermatan sulphate proteoglycans I and II."
RL Biochem. J. 262:823-827(1989).
RN [6]
RP SEQUENCE OF 31-49.
RX MEDLINE=87250639; PubMed=3597437;
RA Fisher L.W., Hawkins G.R., Tuross N., Termine J.D.;

entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```
CC EMBL; AF038127; AAB92652.1; -
CC InterPro; IPR001611; LRR; -
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003592; LRR_out.
DR Pfam; PF00560; LRR; 9.
DR SMART; SM001462; LRRNT; 1.
DR SMART; SM00370; LRR; 3.
DR SMART; SM00013; LRRNT; 1.
KW Glycoprotein; Connective tissue; Extracellular matrix; Proteoglycan;
KW Repeat; Leucine-rich repeat; Signal.
FT SIGNAL 1 16
FT PROPEP 17 30
FT CHAIN 31 360
FT REPEAT 78 99
FT REPEAT 100 123
FT REPEAT 124 146
FT REPEAT 147 168
FT REPEAT 169 194
FT REPEAT 195 218
FT REPEAT 219 239
FT REPEAT 240 263
FT REPEAT 264 286
FT REPEAT 287 309
FT CARBOHYD 34
FT FT
FT CARBOHYD 190 190
FT FT
FT CARBOHYD 212 212
FT CARBOHYD 263 263
FT CARBOHYD 304 304
FT DISULFID 55 68
FT DISULFID 314 347
FT SEQUENCE 360 AA; 39939 MW; 2DAE97CDE16F7C45 CRC64;

Query Match 8.9%; Score 299; DB 1; Length 360;
Best Local Similarity 29.4%; Pred. No. 2.7e-13;
Matches 107; Conservative 47; Mismatches 150; Indels 60; Gaps 13;

QY 15 GPRENRLFDPLLV-----VLLALQLLVAGLVRAQTCPSCVSCSNQFSKVI 61
DQ 17 GPFOQGLDFMLEDASGIGPEDRIHEVLDLEPL-----GPVCPFCQC--HLRVVQ 67
QY 62 CVRKNLREVPDGISTNRLNLHENOIQIKVNSFKHLRHLLEILQLSRNHRTIEICAFN 121
DQ 68 CSDLGLKVPKDLPPDTLLDLQNNKITEIKDGFKNLNLHALLVNNKISKISPGFT 127
QY 122 GLANLNTLELFNDRLTTPNGAFYLSKLKELWLRNPNIESIPSYAFNRIPSLRRLDLG- 180
DQ 128 PLVKLERLYLSKNHLKELPE---KMPKTLQELRVHENEITKVRKAVFNGLNQIMVVELGT 184
QY 181 -ELKRLSYISEGAFGLSNLYLNLAMCNLREI-PNUTPLIKDELDSGNHLSAIRPGS 238
DQ 185 NPLKS-SCIENGAFQGMKLSYIRIADNTITTPGLPP--SLTELHDGNNKITKVDAA 241
QY 239 FQGLHMLQKLMQISOQVTERNAFDNLQSLVEINLAHNNLTLLPHLFTPLHLRIHL 298
DQ 242 LRGLNNLAKGLSFSNISAVDNGSLATPHLRELHNNKLIKVPGL-ADHKYIQVIV 300
QY 299 HNPWNCNDILWLSWIKDKMAPSNTACCARCNTPPNPKGRYIG-----ELDQN 347
DQ 301 HNN--NISA-----VGSNDRCPPCYNTK---KASYSVSLFSPVQVWEIOPS 343
QY 348 YFTC 351
DQ 344 TFRC 347
```

RESULT 13
GPBA_HUMAN

```
ID GPBA_HUMAN STANDARD; PRT; 626 AA.
AC P07359;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Platelet glycoprotein IB alpha chain precursor (GP-IB alpha) (GP1BA)
DE (CD42B-alpha) (CD42B) [Contains: Glycocalicin].
GN GP1BA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=87289655; PubMed=3303030;
RA Lopez J.A., Chung D.W., Fujikawa K., Hagen F.S., Papayannopoulou T.,
RA Roth G.J.;
RT "Cloning of the alpha chain of human platelet glycoprotein Ib: a
RT transmembrane protein with homology to leucine-rich alpha 2-
RT glycoprotein."
RL Proc. Natl. Acad. Sci. U.S.A. 84:5615-5619(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89025874; PubMed=2845978;
RA Wenger R.H., Kieffer N., Wicki A.N., Clemetson K.J.;
RT "Structure of the human blood platelet membrane glycoprotein Ib alpha
RT gene."
RL Biochem. Biophys. Res. Commun. 156:389-395(1988).
RN [3]
RP SEQUENCE OF 17-315.
RX MEDLINE=87289654; PubMed=3497398;
RA Titani K., Takio K., Handa M., Ruggeri Z.M.;
RT "Amino acid sequence of the von Willebrand factor-binding domain of
RT platelet membrane glycoprotein Ib."
RL Proc. Natl. Acad. Sci. U.S.A. 84:5610-5614(1987).
RN [4]
RP DISULFIDE BONDS.
RX MEDLINE=91301149; PubMed=2070794;
RA Hess D., Schaller J., Rickli E.E., Clemetson K.J.;
RT "Identification of the disulphide bonds in human platelet
RT glycocalicin."
RL Eur. J. Biochem. 199:389-393(1991).
RN [5]
RP VARIANT SIBA.
RX MEDLINE=92265982; PubMed=1586750;
RA Murata M., Furihata K., Ishida F., Russell S.R., Ware J.,
RA Ruggeri Z.M.;
RT "Genetic and structural characterization of an amino acid dimorphism
RT in glycoprotein Ib alpha involved in platelet transfusion
RT refractoriness."
RL Blood 79:3086-3090(1992).
RN [6]
RP VARIANT BSS PHE-73.
RX MEDLINE=92110577; PubMed=1730088;
RA Miller J.L., Lyle V.A., Cunningham D.;
RT "Mutation of leucine-57 to phenylalanine in a platelet glycoprotein
RT Ib alpha leucine tandem repeat occurring in patients with an
RT autosomal dominant variant of Bernard-Soulier disease."
RL Blood 79:439-446(1992).
RN [7]
RP VARIANT BSS VAL-172.
RX MEDLINE=93388831; PubMed=7690774;
RA Ware J., Russell S.R., Marchese P., Murata M., Mazzucato M.,
RA de Marco L., Ruggeri Z.M.;
RT "Point mutation in a leucine-rich repeat of platelet glycoprotein Ib
RT alpha resulting in the Bernard-Soulier syndrome."
RL J. Clin. Invest. 92:1213-1220(1993).
RN [8]
RP VARIANT BSS SER-225.
RX MEDLINE=95118882; PubMed=7819107;
RA Simsek S., Norris P., Lozano M., Pico M., von Dem Borne A.E.G.K.,
RA Ribera A., Gallardo D.;
RT "Cys209 Ser mutation in the platelet membrane glycoprotein Ib alpha
```


Db 140 GELQELYLKGNELKTLPLGLLTPPKLEKLSLANNLTLPAGLNGL-----187
Qy 255 IQVTERNAFDNLQSLVEINLAHNLTLLPHDLFTPLHLERHLHHPWNCNCDILWSW 314
Db 188 -----ENLDTLL-----LQENSLYTIKPGFFG-HLLPFAFLHGNPWLNCCEILYFRR 234
Qy 315 WIKDMA-----PSNTACCARCANTPPNLKGRYIG-----ELDONVFTC 351
Db 235 WLODNAENVYVKQGVDMKAMTSNVASVQCDNSDKFPYKYPGKCGPPLGDEGDTLDYD 294
Qy 352 YAPVIVEPADLNTEGMAELKRASTLSLTS-----VSWIT-----PNG 391
Db 295 Y-----PEED-----TEGO-----KVRATRVVVKFTKAHTTWPGLFYSMTASLDSOMPSS 341
Qy 392 TVMTHGAYKVIASVLDGTLNFT-----NVTVDGTGWTVMVNSVGNNTAS-----438
Db 342 LHPTQESTKEOTTPPRKTPNFTLHMSITFSKTPKSTTEPTPS--PTSEPVPEPAPNM 399
Qy 439 ATLNVTAATTTFFSYFSTVVTETMPSODEARTDNNVGPVVDWET-----NVTSL 493
Db 400 TLEPTSPPTTP-----EPTSEPA-PSPTTPEPTPIATISPTILYSATSLI 446
Qy 494 TPQSTRSTEKFTTPVTDINS---GIRGIDEVMKTKIIIG-----CFV 534
Db 447 TP-----KSTFLFTKPVLSLESTKTKTIPELQDPKLVGLQGHLESSRNDPFLHDPDFCCLL 508
Qy 535 AI-----TLMAAVMLVI 546
Db 504 PLGFVVLGLFWLLFASVVLIL 524
RESULT 14
PGS2_CANFA
ID PGS2_CANFA STANDARD; PRT; 360 AA.
AC Q25393;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Bone proteoglycan II precursor (PG-S2) (Decorin).
GN CN OR DCNIC.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Glant T.T.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 244-259 FROM N.A.
RA Venta P.J., Brouillette J.A., Yuzbasliyan-Gurkan V., Brewer G.J.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: BINDS TO TYPE I AND TYPE II COLLAGEN AND AFFECTS THE
CC RATE OF FIBRILS FORMATION. ALSO BINDS TO FIBRONECTIN AND TGF-
CC BETA (BY SIMILARITY).
CC -1- PWM: THE GLYCOSAMINOGLYCAN CHAIN ATTACHED TO DECORIN CAN BE EITHER
CC CHONDROITIN SULFATE OR DERMATAN SULFATE DEPENDING UPON THE
CC TISSUE OF ORIGIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SMALL INTERSTITIAL PROTEOGLYCANS
CC FAMILY.
CC -1- SIMILARITY: CONTAINS 10 LEUCINE-RICH REPEATS (LRR).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U83141; AAB51245.1; -
DR EMBL; L77684; AAA98062.1; -
DR InterPro; IPR001611; LRR.

DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003592; LRR_out.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00560; LRR; 9.
DR Pfam; PF01462; LRRNT; 1.
DR SMART; SM00370; LRR; 2.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_TYP; 1.
KW Glycoprotein; Connective tissue; Extracellular matrix; Proteoglycan;
KW Repeat; Leucine-rich repeat; Signal.
FT SIGNAL 1 16
FT PROPEP 17 30
FT CHAIN 31 360
FT REPEAT 78 99
FT REPEAT 100 123
FT REPEAT 124 146
FT REPEAT 147 168
FT REPEAT 169 194
FT REPEAT 195 218
FT REPEAT 219 239
FT REPEAT 240 263
FT REPEAT 264 286
FT REPEAT 287 309
FT CARBOHYD 34 34
FT CARBOHYD 212 212
FT CARBOHYD 263 263
FT CARBOHYD 304 304
FT DISULFID 55 68
FT DISULFID 314 347
SQ SEQUENCE 360 AA; 99880 MW; 99881A9C812906 CRC64;
Query Match 8.8; Score 295; DB 1; Length 360;
Best Local Similarity 29.2; Pred. No. 5e-13;
Matches 106; Conservative 45; Mismatches 154; Indels 58; Gaps 12;
Qy 15 GPFRNALFDPLL-----VVLALQLLVAGLVRAQTCPSCVCSNQFSKVI 61
Db 17 GPFQRLGDFMLEASGIGPEDRAPDMPDLELL-----GPVCPFCQC--HLRVVQ 67
Qy 62 CVRNKLRVDPGISTNTRLLNHNQIIVKNSFKHLRHLRIQLSRNHTTEIGAFN 121
Db 68 CSDGLDKVPKDPDPTDILLQNNKITEIKDGFKNLKNLHTLLVNNKISKISPGFT 127
Qy 122 GLANLTLELFDNRLTTIPNGAFVLSKLKELLRNPNIESIPSVAFNRIPSLRLDAG- 180
Db 128 PLLKLERLYLSKNHLKELPE---KMPKTLQELRAHENEITVKRAVFNGLNQMVIVELGT 184
Qy 181 -ELKRLSYISEGAPEGLSNRLNLCNREIPNLTPLIKLELDELQSLGNSHLSAIRPGSF 239
Db 185 NPLKS-SGIENGAFQGMKLSYIRIADNTITTIPOGLP-PSLTELHLEGNKITKVDASSL 242
Qy 240 QGLMHLQKLMQIQIOVIERNAFDNQSLVEINLAHNLTLLPHDLFTPLHLERHLH 299
Db 243 KGLNLAKLGLSFSISAVDNGTLANTPLHRELHLDNKNKIRVPGGL-AEHKIQVYVLH 301
Qy 300 HNPWNCIDLWLSWIKDMAPSNTACCARTPPNLKGRYIG-----ELDONV 348
Db 302 NN--NISA-----VGSNDFCPGYNK---KASISGVSLSPNPVQYWEIQPST 344
Qy 349 FTC 351
Db 345 FRC 347
RESULT 15
PGS2_PIG
ID PGS2_PIG STANDARD; PRT; 360 AA.
AC Q9XSD9; Q9XSH4;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Bone proteoglycan II precursor (PG-S2) (Decorin).
GN DCN.
OS Sus scrofa (Pig).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID:9822;
RN [1]
RP SEQUENCE FROM N.A. (LONG FORM).
RC STRAIN=YORKSHIRE;
RA Stephenson S., Schnoke M., Vesely I.;
RT "Cloning of the porcine decorin gene";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A. (SHORT FORM).
RC STRAIN=YORKSHIRE; TISSUE=Aorta;
RA Stephenson S., Schnoke M., Vesely I.;
RT "Alternatively spliced version of the porcine decorin gene";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: BINDS TO TYPE I AND TYPE II COLLAGEN AND AFFECTS THE
CC RATE OF FIBRIL FORMATION. ALSO BINDS TO FIBRONECTIN AND TGF-
CC BETA (BY SIMILARITY).
CC -1- PPM: THE GLYCOSAMINOGLYCAN CHAIN ATTACHED TO DECORIN CAN BE EITHER
CC CHONDROITIN SULFATE OR DERMATAN SULFATE DEPENDING UPON THE
CC TISSUE OF ORIGIN.
CC -1- SIMILARITY: BELONGS TO THE SMALL INTERSTITIAL PROTEOGLYCANS
CC FAMILY.
CC -1- SIMILARITY: CONTAINS 10 LEUCINE-RICH REPEATS (LRR).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF125537; AAD23578.1; -;
DR EMBL; AF140270; AAD33862.1; -;
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003592; LRR_Out.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00560; LRR_9.
DR Pfam; PF01462; LRRNT; 1.
DR SMART; SM00370; LRR; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_Typ; 2.
KW Glycoprotein; Connective tissue; Extracellular matrix; Proteoglycan;
KW Repeat; Leucine-rich repeat; Signal; Alternative splicing.
FT SIGNAL 1 16
FT PROPEP 17 30
FT CHAIN 31 360 BONE PROTEOGLYCAN II.
FT REPEAT 78 99 LRR 1.
FT REPEAT 100 123 LRR 2.
FT REPEAT 124 146 LRR 3.
FT REPEAT 147 168 LRR 4.
FT REPEAT 169 194 LRR 5.
FT REPEAT 195 218 LRR 6.
FT REPEAT 219 239 LRR 7.
FT REPEAT 240 263 LRR 8.
FT REPEAT 264 286 LRR 9.
FT REPEAT 287 309 LRR 10.
FT CARBOHYD 34 O-LINKED (GLYCOSAMINOGLYCAN) (BY
FT SIMILARITY).
FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 55 68 BY SIMILARITY.
FT DISULFID 314 347 POTENTIAL.
FT VARSPPLIC 281 318 MISSING (IN SHORT ISOFORM).
SQ SEQUENCE 360 AA; 39899 MW; 8573DE8DEBA7509 CRC64;

Query Match 8.8%; Score 295; DB 1; Length 360;
Best Local Similarity 30.5%; Pred. No. 5e-13;
Matches 97; Conservative 44; Mismatches 139; Indels 38; Gaps 10;
QY 47 CPSVCSGNSQFSKVICVRKNLRVDPGISTNRLNLHENOIOIKVNSFKHLRHLIQ 106
DB 55 CPFRQCQ--HLRVVQCSDLGLDKVPKLPDPDTALLDQNNKITEIKDGFKNLNLHTLI 112
QY 107 LSRNHRTIETGAENGLANLNTLELNFONRLTTPNGAFVYLSKLKLWLRNNTIESIPSY 166
DB 113 LINNKISKISPGAFAPLVKLERLYLSKNQKELPE---KMPKTIQELRVHENEITKVRKA 169
QY 167 AFNRIPSLRRDLG--ELKRLSYISEGAFGLSNRLYLNLAMCNLRLEIPNLTPIIKLDEL 224
DB 170 VFNGLNQIMVVELGTNPLKS-SGIENGAFQGMKKLSYIRIADTNTITTPQGLP-PSLTEL 227
QY 225 DLSGNHLSAIRPGSFOGLMHLOKLMWIOSQIOVIERNAFDNLQSLVEINLAHNNLTLLPH 284
DB 228 HLDGNKISKYDAASLKLGLNNLAKLGLGFNSISTVDNGLANTPHLRHLNKNLKNKVP 287
QY 285 DLFTPLHHLERIHLLHNPWNCNCDILWSWIKDKMAPSNTACCARCTPPNPKRGYIG-- 342
DB 288 GL-AEHKYYIQVYVYLHN--NISA-----VGSNDFCPCPYNTK---KASYSGVS 329
QY 343 -----ELDQNYFTC 351
DB 330 LFSNPVQYWEIQPSTERC 347

Search completed: August 26, 2002, 15:45:35
Job time: 247 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 15:41:03 ; Search time 34.22 Seconds
(without alignments)
3235.440 Million cell updates/sec

Title: US-09-905-056-292
Perfect score: 3362
Sequence: 1 MLNKMTHLHPOQIMIGPRNR.....VHEPLLIRMSKDNVQETQI 640

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 17994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPTEMBL_19.*

2: sp_archaea.*

3: sp_bacteria.*

4: sp_fungi.*

5: sp_human.*

6: sp_invertebrate.*

7: sp_mammal.*

8: sp_mhc.*

9: sp_organelle.*

10: sp_phase.*

11: sp_plant.*

12: sp_rodent.*

13: sp_virus.*

14: sp_vertebrate.*

15: sp_unclassified.*

16: sp_rvivirus.*

17: sp_bacteriap.*

18: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3362	100.0	640	Q9HCJ2	Q9hcj2 homo sapien
2	1865	55.5	653	Q9HBW1	Q9hbw1 homo sapien
3	1840	54.7	649	Q96A85	Q96a85 homo sapien
4	1376.5	40.9	422	Q9NT99	Q9nt99 homo sapien
5	1153	34.3	441	Q99PH1	Q99ph1 mus musculus
6	583.5	17.4	614	Q96FE5	Q96fe5 homo sapien
7	578.5	17.2	614	Q9N008	Q9n008 macaca fasc
8	577.5	17.2	614	Q9D1T0	Q9d1t0 mus musculus
9	575	17.1	606	Q9BZ20	Q9bz20 homo sapien
10	480.5	14.3	719	Q96N16	Q96ni6 homo sapien
11	466	13.9	1021	Q9V430	Q9v430 drosophila
12	442.5	13.2	832	Q9ULH4	Q9ulh4 homo sapien
13	437.5	13.0	789	Q9BE71	Q9be71 macaca fasc
14	432	12.8	705	O43377	O43377 homo sapien
15	432	12.8	708	Q9H3W5	Q9h3w5 homo sapien
16	431.5	12.8	492	Q99KT6	Q99kt6 mus musculus

ALIGNMENTS

RESULT 1

Q9HCJ2 ID Q9HCJ2 PRELIMINARY; PRT; 640 AA.

AC Q9HCJ2; 01-MAR-2001 (TrEMBLrel. 16, Created)

DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE KIAA1580 PROTEIN (FRAGMENT).

GN KIAA1580.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20450683; PubMed=10997877;

RA Nagase T., Kikuno R., Nakayama M., Hirotsawa M., Ohara O.;

RT "Prediction of the coding sequences of unidentified human genes.

RT XVIII. The complete sequences of 100 new cDNA clones from brain which

code for large proteins in vitro."

RL DNA Res. 7:273-281(2000).

DR EMBL; AB046800; BAB13406.1; .

DR HSSP; P22888; 11UT.

DR InterPro; IPR003599; Iq.

DR InterPro; IPR003598; Iq_c2.

DR InterPro; IPR003600; Iq_like.

DR InterPro; IPR003006; Iq_MHC.

DR InterPro; IPR001611; LRR.

DR InterPro; IPR000483; LRR_Cterm.

DR InterPro; IPR000372; LRR_Nterm.

DR InterPro; IPR003592; LRR_out.

DR InterPro; IPR003591; LRR_typ.

DR Pfam; PF00047; Iq; 1.

DR Pfam; PF00560; LRR; 9.

DR Pfam; PF01463; LRRCT; 1.

DR Pfam; PF01462; LRRNT; 1.

DR PRINTS; PR00019; LEURICHRPT.

DR SMART; SM00409; IG; 1.

DR SMART; SM00408; IGC2; 1.

Q9p231 homo sapien
Q73675 xenopus lae
Q9cyk3 mus musculu
Q61809 mus musculu
Q96ja1 homo sapien
Q9by88 homo sapien
Q97860 mus musculu
Q96rj5 homo sapien
P70193 mus musculu
O15335 homo sapien
Q9esy6 rattus norv
Q9nzul homo sapien
Q9wvb5 mus musculu
Q94898 homo sapien
Q88279 rattus norv
Q95710 homo sapien
Q94813 homo sapien
Q92626 homo sapien
Q95226 mus musculu
Q9y5q7 homo sapien
Q9rlb9 mus musculu
Q9wug5 rattus norv
Q70210 rattus norv
Q9de37 brachydanio
Q90wz3 xenopus lae
Q9wvc1 rattus norv
Q9nzu0 homo sapien
Q9p259 homo sapien
Q75093 homo sapien

17 430 12.8 730 4 Q9P231
18 426.5 12.7 718 13 Q73675
19 426.5 12.7 788 11 Q9CYK3
20 423.5 12.6 716 11 Q61809
21 420 12.5 1093 4 Q96JA1
22 420 12.5 1094 4 Q9BY88
23 419 12.5 707 11 Q97860
24 418 12.4 359 4 Q96RJ5
25 417.5 12.4 1091 11 P70193
26 414 12.3 359 4 O15335
27 406 12.1 707 11 Q9ESY6
28 401 11.9 674 4 Q9NZUL
29 401 11.9 1531 11 Q9WVB5
30 399.5 11.9 1065 4 Q94898
31 399 11.9 1531 11 Q88279
32 397.5 11.8 1521 4 Q95710
33 397.5 11.8 1529 4 Q94813
34 397 11.8 1496 4 Q92626
35 395.5 11.8 358 11 Q95226
36 395.5 11.8 1525 4 Q9Y5Q7
37 394.5 11.7 1521 11 Q9RLB9
38 394.5 11.7 1530 11 Q9WUG5
39 393 11.7 358 11 Q70210
40 389.5 11.6 1515 13 Q9DE37
41 387.5 11.5 1530 13 Q90WZ3
42 386.5 11.5 796 11 Q9WVC1
43 385 11.5 649 4 Q9NZU0
44 385 11.5 662 4 Q9P259
45 385 11.5 1534 4 Q75093

```
DR SMART; SM00410; IG_like; 1.
DR SMART; SM00370; LRR; 6.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_TYP; 8.
KW Immunoglobulin domain.
FT NON_TER 1
SQ SEQUENCE 640 AA; 71949 MW; 6058974872636838 CRC64;

Query Match 100.0%; Score 3362; DB 4; Length 640;
Best Local Similarity 100.0%; Pred. No. 2.3e-240;
Matches 640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLNKMTHLPQOIMIGPFRNALFDPLLVLLALQLLVVAGLVRAQTCPVSCSCSNQFSKV 60
Db 1 MLNKMTHLPQOIMIGPFRNALFDPLLVLLALQLLVVAGLVRAQTCPVSCSCSNQFSKV 60

Qy 61 ICVRKNLREVDPGISTNTRLLNHNENQIIKVNSEFKHLRHLLEILQLSRNHRTIEIGAF 120
Db 61 ICVRKNLREVDPGISTNTRLLNHNENQIIKVNSEFKHLRHLLEILQLSRNHRTIEIGAF 120

Qy 121 NGLANLNTLEFDRNLRTIPNGAFVYLSKELKELWLRNNPIESIPSYAFNRIPSLRRDLG 180
Db 121 NGLANLNTLEFDRNLRTIPNGAFVYLSKELKELWLRNNPIESIPSYAFNRIPSLRRDLG 180

Qy 181 ELKRLSYSEGAFGLSLNRYLNLMCNLRIPNLTPLIKLDELDSLGNHLSAIRPGSFQ 240
Db 181 ELKRLSYSEGAFGLSLNRYLNLMCNLRIPNLTPLIKLDELDSLGNHLSAIRPGSFQ 240

Qy 241 GLMHQLKWMTQSQIOVIERNAFNLQSLVEINLAHNNTLLPHDLFTPLHHLERHLLH 300
Db 241 GLMHQLKWMTQSQIOVIERNAFNLQSLVEINLAHNNTLLPHDLFTPLHHLERHLLH 300

Qy 301 NPWNCNDILWLSWIKMDSNTACCARNTPNLRKGYIGELDONFTCYAPVIVEPP 360
Db 301 NPWNCNDILWLSWIKMDSNTACCARNTPNLRKGYIGELDONFTCYAPVIVEPP 360

Qy 361 ADLVNTEGMAELKCRASTSLTSYTSWITPNTGVTMGAYKRYAVLSLDTLNFNTVQD 420
Db 361 ADLVNTEGMAELKCRASTSLTSYTSWITPNTGVTMGAYKRYAVLSLDTLNFNTVQD 420

Qy 421 TGMVTCMVNSVGNVTASATLNVTAAATTPSYSTVTVETMPSQDEARTDNNVGTP 480
Db 421 TGMVTCMVNSVGNVTASATLNVTAAATTPSYSTVTVETMPSQDEARTDNNVGTP 480

Qy 481 VVDNETNVTSLTPQSTRTEKFTTIPVTDINSIGIPGIDEVMTKTIIGCFVAITLMA 540
Db 481 VVDNETNVTSLTPQSTRTEKFTTIPVTDINSIGIPGIDEVMTKTIIGCFVAITLMA 540

Qy 541 AVMLVIFYKMRKHRRQNHHAAPTVEIINVDDITGDTMESHLPMPAIEHEHLNHYNS 600
Db 541 AVMLVIFYKMRKHRRQNHHAAPTVEIINVDDITGDTMESHLPMPAIEHEHLNHYNS 600

Qy 601 YKSPFNHTTNTVNTINSIHSSVHEPPLLIRMSKDNVQETQI 640
Db 601 YKSPFNHTTNTVNTINSIHSSVHEPPLLIRMSKDNVQETQI 640
```

RESULT 2

```
Q9HBW1 PRELIMINARY; PRT; 653 AA.
AC Q9HBW1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE BRAIN TUMOR ASSOCIATED PROTEIN NAG14.
GN NAG14.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RA Wang J., Bin L., Jiang N., Li G.;
RT "Homo sapiens brain-specific gene (BAG), downregulated in brain tumor, mRNA.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF196976; AAC28019.2; -
DR HSP; P23945; 1XUN.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003592; LRR_out.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00560; LRR; 8.
DR Pfam; PF01463; LRRCT; 1.
DR Pfam; PF01462; LRRNT; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00408; IGc2; 1.
DR SMART; SM00410; IG_like; 1.
DR SMART; SM00370; LRR; 3.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_TYP; 7.
KW Immunoglobulin domain.
SQ SEQUENCE 653 AA; 72717 MW; 38159C81F6850E37 CRC64;

Query Match 55.5%; Score 1865; DB 4; Length 653;
Best Local Similarity 56.4%; Pred. No. 1.2e-129;
Matches 363; Conservative 107; Mismatches 130; Indels 44; Gaps 14;

Qy 29 VLLALQLLVAGLVRA-----QTCPSVCSNQSFKVICVRKNLREVDPGISTNTRLLN 82
Db 22 VYLAQVWILCAATAAASAGPQNCPSVCSNQSFKVYCTRRGLSEVPQGPSNTRYLN 81

Qy 83 LHENOIQIIVKNSFKHLRLEILOSRNHRTIEIGAFNGLANLNTLELFDNRLTTPNG 142
Db 82 LMENNIOIADTFRHLHLEVLQGRNSIRQIEVGAENGLASLNTLELFDNRLTTPNG 141

Qy 143 AFVLSKLKELWLRNNPIESIPSYAFNRIPSLRRDLGELKRLSYISEGAFGLSNRYL 202
Db 142 AFEVLSKLRLWLRNNPIESIPSYAFNRVPSLMRLDLGELKKLEYISEGAFGLNRYL 201

Qy 203 NLAMCNLRIPNLTPLIKLDELDSLGNHLSAIRPGSFOGLMHLOKLMWISQIOVIENRA 262
Db 202 NLGMCNKDMPNLTPLVGLLEELMSGNHPFPIRPGSFHGLSLKKLWYMNQVSLIERNA 261

Qy 263 FDNLOSLVEINLAHNNTLLPHDLFTPLHHLERTHLHNPNWNCNDILWLSWIKMDAPS 322
Db 262 FDGLASVELNLAHNNTLLPHDLFTPLRYLVELHLHNPNWNCNDILWLAWLREYIPT 321

Qy 323 NTACCARNCTPPNLKGRYIGELDONFTCYAPVIVEPPADLVNTEGMAELKCRASTSLT 382
Db 322 NSTCCGRCHAPMHRGRYLVVEVDQASFCQSAPPFIMDAPDLNISBGRMAELKCR-TPMS 380

Qy 383 SVSWITPNTGVTMGAYKRYAVLSLDTLNFNTVQDTGTMVTCMVNSVGNVTASATLN 442
Db 381 SVKLLPNTGTVLSHARHPRIKSVLNDGTLNFSHVLSDTGYVTCVMTNVAGNSASAYLN 440

Qy 443 VTAA--TTTPFSYFTVTVETMPSQDEARTDNNVGTPPVVDNETNV-----TTSLT 494
Db 441 VSTAEIANTSYSTFTVTVETTEISPED---TTRKYKPPV-----TTSQGPAYTSTTT 492

Qy 495 P--QSTRSTEXTFTIPVTDINSIGI-PGIDEVMKTKIIGCFVAITLMAAVLVIFYKMR 551
Db 493 VLIQTR-VPKQVAVPATDITDKQTSLEVMKTKIIGCFVAVTLAAAMLVIFYKLR 551

Qy 552 KOHRQNHHAAPTVEIINVDDITGDTPM-----ESHLPMPAIEHEHLNHYNS 600
Db 552 KOHRQNHHAAPTVEIINVDDITGDTPM-----ESHLPMPAIEHEHLNHYNS 600
```



```

Db 540 ATV-PEP-FDIKTLIIATMGFIS-----F 562
QY 534 VAITLMAAVMLVIFY--KMRKHQRHONHAPTRVELINVDDEITGTPMESHLP 587
Db 563 LGVVLFLCLVLLFLWSRGKGNTHKNIETIYVPRKSDAGIS-----SADAPRKFNMKM 613

RESULT 7
Q9N008 ID Q9N008 PRELIMINARY; PRT; 614 AA.
AC Q9N008;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE HYPOTHETICAL 69.2 KDA PROTEIN.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=CEREBELLUM CORTEX;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB046639; BAB03557.1; -
DR HSP; P23945; Ixun.
DR InterPro; IPR003598; Ig_C2.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003592; LRR_Out.
DR InterPro; IPR003591; LRR_Typ.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00560; LRR; 9.
DR Pfam; PF01463; LRRCT; 1.
DR Pfam; PF01462; LRRNT; 1.
DR Pfam; PF00408; IGC2; 1.
DR SMART; SM00370; LRR; 5.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_Typ; 1.
KW Hypothetical protein; Immunoglobulin domain.
SQ SEQUENCE 614 AA; 69187 MW; BA6C8BC7C993BE9A CRC64;

Query Match 17.2%; Score 578.5; DB 6; Length 614;
Best Local Similarity 27.6%; Pred. No. 1.6e-34;
Matches 181; Conservative 96; Mismatches 242; Indels 137; Gaps 16;

QY 20 RALEPDLVLLALQLLVAGLV--RAQTCPSCVCSNQFSKIVCRKNLRVDPDGISTN 77
Db 7 RMSPEPLACQWPIILLVLGSLSGATGCPPECECAQDRAVLCHKRKFVAVPEGIPT 66
QY 78 TRLLNLHENOIIKVNFSFKHLRHLLEIQLSRNHIRFIEIGAFNGLANLWLEFDNRLT 137
Db 67 TRLLDLGKNRIKTLNQDEFASFPHLEELNENITVSAVEFGAFNNLFLNRLTLGRSLRK 126
QY 138 TIPNGAFVYLSKELKELWLNRPNIPSPYAFNRIPSLRRLLDLGLKRLYSIEGAFEGLS 197
Db 127 LIPGVGTGLSNLTKLDISENKIVILLDYMFQDLYNLSKLEVG-DNLVYISHRAFGSLN 185
QY 198 NLRVNLAMCNLRIP----- 213
Db 186 SLEQLTLEKCNLTSPITSEALSHLGLVLRHLNLINAIRDYSFKRLYRLKVLISHWPY 245
QY 214 -----NLT-----LIKLDLDSGNHLSAIRPGSFGLM 243

```

```

Db 246 LDTMTPNCLYGLNLTSLSITHCNLTAVPYLAVRHLVYLRFLNLNLSYNPISTIEGSMHELL 305
QY 244 HLKRLWIOSIQIVERNAPDNLOSILVEINLAHNNLTLPHDLTPLHLHRIHLHNPW 303
Db 306 RLQEIQLVGGOLAMVEPYAFRLNLYLRVLNVSNGNLTLEESVHSGNLTLLDSNPL 365
QY 304 NCNCIDIWL---SWWIKDMAPSNACCARCNTPPNLKAGRYIGE---LDQNYFTCYADVI 356
Db 366 ADCRLLWVFRRLNFRNQPT-----CATPEVQGEKFPDVLNPNYFCRRARI 420
QY 357 VEPPA-DLNVTEGMAELKCR-A-STSLTSVSWITPNTGTHGAYKVRIAVLSOGTLNFT 414
Db 421 RDRKAQVFEVDEGHTVQFCRADGDPPIAILWSPRKLVS-AKSNGLRTVFPDCTLEVR 479
QY 415 NTVQDTGMVTCMVNSVGNVTASATLNVT-AATTPEYSYFTVVTETMPSQDEARTTD 473
Db 480 YAVQVDNGTYLCIAANAGGDSMPAHLHVRYSIDPWPHQPNKTFATFPNQGEGEANSTR 539
QY 474 NNVGTPVVDWETTNVTSLTPQSTRTKFTTIPVTDINSIGPIDEVMKTKIIIGCF 533
Db 540 ATV-PEP-FDIKTLIIATMGFIS-----F 562
QY 534 VAITLMAAVMLVIFY--KMRKHQRHONHAPTRVELINVDDEITGTPMESHLP 587
Db 563 LGVVLFLCLVLLFLWSRGKGNTHKNIETIYVPRKSDAGIS-----SADAPRKFNMKM 613

RESULT 8
Q9D1T0 ID Q9D1T0 PRELIMINARY; PRT; 614 AA.
AC Q9D1T0;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE ADULT MALE TESTIS CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,
DE CLONE:4930471K13, FULL INSERT SEQUENCE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TESTIS;
RX MEDLINE=21085660; PubMed=11217851;
RA Akawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Salto T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Straubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001)
DR EMBL; AK027262; BAB32403.1; -
DR HSP; P23945; Ixun.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003598; Ig_C2.
DR InterPro; IPR003600; Ig_Like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.

```



```
Db 311 TIEPHSFOGLRFLRVNLVNSQNLLETLEBNVSPRALEVLNINNPLACDCLRLWLQRO 370
Qy 313 -SWIKMAPNTACCARCNTPPNLKGRYIGELDON-----YTCYAPVIVPPAD-LNVT 366
Db 371 PTLQGGQOP-----MCAGPDITERSPKDFHSTALSFFYTCCKPKIKREKKLQHLVD 423
Qy 367 EGMAELKCRASLTSTSV-SWITNG---TWMTGAYKVRZAVLSDGTPLNTNTVQDGT 422
Db 424 EGQVQLECSADGDPQPVISWVTPRRRFTTKSNG---RATVLGDDGTLEIRFAQDDSG 479
Qy 423 MYTCWNSVGNNTASATLNVTAATTPFSYFSTVVTETMEPSQDEARTDNNVGPV 482
Db 480 MYVCIASNAAGNDFTASLTGKGFASDRFLYANRTPMTDNDSTISNGTNAN---TFSL 536
Qy 483 DWETNTVTSLTPOSTRSTKFTFTIPVTDINSIGIPIDEVMKTKIIIGCFVAITLMAAV 542
Db 537 DLKTLVSTAM-----GCTFLGVVLCF 559
Qy 543 MLVIFYKMR-KQHRON---HHAPTRVVEIINVDDEITG 577
Db 560 FLLLFVWSRGKHKHNSIDLEYVPRKNGAV-VEGEVAG 597

RESULT 10
Q96NI6
ID Q96NI6 PRELIMINARY; PRT; 719 AA.
AC Q96NI6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CDNA FLJ30803 FIS, CLONE FEBRA2001245, WEAKLY SIMILAR TO NAG14.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA Ninomiya K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
RA Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K.,
RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K.,
RA Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK055365; BAB70910.1;
SQ SEQUENCE 719 AA; 79472 MW; CB301F219705FF4B CRC64;

Query Match 14.3%; Score 480.5; DB 4; Length 719;
Best Local Similarity 25.58; Pred. No. 3.5e-27;
Matches 169; Conservative 78; Mismatches 249; Indels 167; Gaps 16;

Qy 29 VLLALQLLVAGLVRAQTCPSVCS--NPFSKVICVRKNLREVPDGTSTNRLNLNHN 86
Db 4 ILFVFLGIA--VKAQICPRVCQILSPNLATLCAKGLLFVPPNDRITVELRLADN 61
Qy 87 QIQIKVNSFKLRLHLQLSRNHRTIEIGANGNLANTLELFDNRLTTPNGAPVY 146
Db 62 FVTNKKRDFANMTSLVDLTSRNTISPTIPAFADLRNLRALHNSNRLTKITNMFSG 121
Qy 147 LSKLKLWLNRNPSTPSYAFNRIPSLRRLDGLGELKRLSYISEGAFGLSNRLYNLAM 206
Db 122 LSNLHLILNNQTLISSTAFDDV----- 146
Qy 207 CNLREIPNLPLIKDELDSNHLSAIRPGSFGQLMHLQKLWMTQSOIQUIERNAFDNL 266
Db 147 -----FALEELDSLNNLETIPWDAVEKMSVLSHTLSLDHNNIDNIPKGTSHL 194
Qy 267 QSLVEINLAHNNLPLPHDLTPLHLERI-----HLHNPWNCNCDILWL 312
```

```
Db 195 HKMTRLDVTSNKLOKLPD---PLFORAQVLATSGIISPSTFALSFGNPLNCCELLWL 251
Qy 313 SWIKIDMAPNTACCARCNTPPNLKGRYIGELDONFYTCYAPVIVPPADLNVTGMAE 372
Db 252 -----RLSRDDLETCASPLLTGRYFWSIPEEFCEPPLITRHTHEMRVLEQGRAT 305
Qy 373 LKCRATSTLSVSWITPMTGTHGAYKVRZAVLSDGTPLNTNTVQDGTMYTCWNS 431
Db 306 LRCRARGDPEPAIHWSPEGLISNA---TRSLVVDNGTLDLITVTKDTGAFTCIASNP 362
Qy 432 VGNTTASA-----TLNVTAATTPFSYFSTVVTETMEPSQDEARTDNNVGPV 482
Db 363 AGEATQIVDLHIIKLPHLLNSTNIHEPDGSSDSTSTKSGNSTSSNGDTKLSDQKIV 422
Qy 483 DWETNTVTSLTPOSTRSTKFTFTIPVTDINSIGIPIDEVMKTKIIIGCFVAITLMAAV 542
Db 423 VAEATSTALLKFNFORNIPGIRMFQIOYNGTYDDTLVYRMIPPTSKFL--VNNLAGT 480
Qy 517 -----PGIDEVMKT-----TKIII--GCFVAI 536
Db 481 MYDLCVLAIDGDTSLTATRVVGCIOFTTBQDYVRCHFMOSQFLGGTMIIGIIVAS 540
Qy 537 TLMRAVMLVIFYKMRKQHRONHHAPTRVVEIINVDDEITGTPMESHLPMP-----AI 590
Db 541 VLVFIILMIRYK-----CNNNGQHKVTKVSNVYSQTNGAQIQGCSVTLPQSVSKOAV 594
Qy 591 EHE 593
Db 595 GHE 597

RESULT 11
Q9V430
ID Q9V430 PRELIMINARY; PRT; 1021 AA.
AC Q9V430;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE KEK3 PROTEIN.
GN KEK3 OR BG:DS04862.1 OR CG4192.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balley R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burkova K., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
```


Db 304 AVDNLPLRTEATNPNRLSYIHPNAPFLPKLESMLNSALSALYHGHTIESLPNLKEI 363
Qy 297 HLHNPWNCIDILWSW---WIDMAPSNACCARCNTPPNLKGRYIGELD-QNYETCY 352
Db 364 SIHNPTRCDDVIRWMNMKNKNIIRFMEPDSLFCV-----DPPEFOGQNRQVHFDRMMEIC 419
Qy 353 APVTVEP---PADLNVTEGMAAELKCRASTS-LTSVSWITPNGTVMTHGAYKRVIAVLSDG 409
Db 420 LPLAPESFPNVLNVEAGSYVFCRATAEPOPEIYWIITPSGOKLLPNTLTKDYVHSEG 479
Qy 410 TLNFTNTVQDGYMTCMVSNVGNVTASATFLNVTAAATTPFSYFSTVVTMETPESQDEA 469
Db 480 TLDINGVTPKEGGLYTCIATNLVGLADLKSVMKV-----DGS 516
Qy 470 RTTNNVGPVVDWEITNVTTSLTPQSTRSTKFTPIPTVDINSIGPIDEVMTTKII 529
Db 517 FPQDNN-----GSLNKKIRDIQA-----NSVLVSWKA- 543
Qy 530 IGCFAVITLMAAVLVIFPKMKQHRRHONHAPRTVEIINVDDIEITGDTPMESHLPMPA 589
Db 544 ---SSKILSSVKWTAFTVKTENSAAQSARIPS-DVKVYNL-THLNPSTEYKICIDIPT 597
Qy 590 IEHEHLN-----HNSYKSPFNHTTV 611
Db 598 IYQNRKRCVNVTTKGLHPDQKEYEKNNTTL 629

RESULT 2

US-08-986-485-5
; Sequence 5, Application US/08986485
; Patent No. 6046030
; GENERAL INFORMATION:
; APPLICANT: WU, SHUJIAN
; APPLICANT: SWEET, RAYMOND
; APPLICANT: TRUNEH, ALEMSEGED
; TITLE OF INVENTION: A HUMAN LIG-1 HOMOLOG (HLIG-1)
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/986,485
; FILING DATE: 08-DEC-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/059,448
; FILING DATE: 22-SEP-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70264
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1091 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-986-485-5

Query Match 12.4%; Score 417.5; DB 3; Length 1091;
Best Local Similarity 23.1%; Pred No. 5.3e-27;
Matches 156; Conservative 87; Mismatches 208; Indels 223; Gaps 19;
Qy 81 LNLHENOQIIKIVNSFKHLRHLLEILQLSRNHRITIEIGAFNGLANLNTLELFDNRLTTP 140
Db 218 LDLNRRNRIIEGLTFCGLDSLEVLRLQRNNISRLTDCAFWGLSKMHVHLEYSNLSVEV 277
Qy 141 NGAFVYLSKLKELWLRNPNIESIPS-----YAFNRIPSLRLDLGELKRL-- 185
Db 278 SGSLYGLTALHQLHLNNSISRIQRDGSFCQKLHELILSFNNLRLDEESLAELSLSI 337
Qy 186 -----SYISGAFEGLSNLRVNLNLMCNLRPEINLTPLIKDELDSLNSHLSATRPG 237
Db 338 LRLSHNAISHAEGAFKGLSLRVL-----LDHNEISTIEDT--SG 378
Qy 238 SFQGLMHLQKLMWIOSQIVIERNAFDNLQSLVINLAHNNLTLLPHDLTPLHHLERIH 297
Db 379 AFTGLDNLKSLTLFGNKIKSVAKRAFGLESLEHLNLGENAIRSVQFADFAMKMLKELY 438
Qy 298 LHHNPWNCIDILWSWI---KDMAPSNACCARCNTPPNLKGRYI-GELQNY----- 348
Db 439 ISSSEFLCDCQLKWLPPMLGMRMLQAFVTATCAH---PESLKGOSIFSVPDSFVCDPFP 495
Qy 349 -----FTCYA----- 353
Db 496 KPQIITQPTTMVVGVKDIRFTCSAASSSSPMTFAMKKDNEVLNADNMFVHRAQDG 555
Qy 354 -----PVIVEPPADLNVTEG 368
Db 556 EVMEYTTILHLRHVTEGHEGRYQCIITNHFGSTYSHKARLVNVLPSFTKIPHDIAIRTG 615
Qy 369 MAELKCRASLT-SVSWITPNGTVMTHGAYKRVIAVLSDGTLNF-TNVTQDGTGMYTC 426
Db 616 TTRLECAATGHPNPQIAWQKDGDTDF-AARERMMHVMPPDDVFFITDVKIDDMGVYSC 674
Qy 427 MVSNSVGNNTASATLNVTAATTPFSYFSTVVTMETPESQDEARTDNNV-----GP 478
Db 675 TAQNSAGSVSANATLVLE-----TPSLAVPLEDRVVTGTVGTFVAFQCKATGSP 722
Qy 479 TPVVDWETNTVTTSLT-----PQSTRSTEKTTI 507
Db 723 TPRTWLKGGRRPLSLTERHHFTPGNQLLVQNVMIIDDAGRYTCMSNPLGTERAHSQLSI 782
Qy 508 PVTDINSIGPIDEVMTTKII-IGCFVAITLMAAVMLVIFYKMKRKHRRHONHAPRTTV 566
Db 783 LPT-----PCRKDGTIVGIFTIAVVCISIVLTSLVWVCIYQTRK---KSEYSVTWTD 833
Qy 567 EIIINVDDIEITGDT 580
Db 834 ETI-----VPPDVP 842

RESULT 3

US-08-986-485-2
; Sequence 2, Application US/08986485
; Patent No. 6046030
; GENERAL INFORMATION:
; APPLICANT: WU, SHUJIAN
; APPLICANT: SWEET, RAYMOND
; APPLICANT: TRUNEH, ALEMSEGED
; TITLE OF INVENTION: A HUMAN LIG-1 HOMOLOG (HLIG-1)
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA: US/08/986,485
APPLICATION NUMBER: US/08/986,485
FILING DATE: 08-DEC-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/059,448
FILING DATE: 22-SEP-1997
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70264
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1101 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
us-09-986-485-2

Query Match 11.98; Score 401; DB 3; Length 1101;
Best Local Similarity 22.08; Pred. No. 1.4e-25;
Matches 172; Conservative 111; Mismatches 265; Indels 234; Gaps 27;
Qy 26 LVLVLLAQLLVAGLVRAQTCPSCVCSNQ-----FSKVICVRKNLREVDPGISTNRL 80
Db 21 LWLILLRLEPYTAAGPRA-PCAACTCAGDPTCAGDSLCCGGRLAALPGDLPSTWRS 79
Qy 81 LNL-----HE----- 85
Db 80 LNLSYNKLAEIDPAGFEDLPNLQEVYLNHHELTAVASLGAGSSOVVALFLOQQQNRSLDG 139
Qy 86 -----NQIILKIVNSFKHLRHLLETLQLSRNHIETIGAFNGLA-NLNT 128
Db 140 SOLKAYLSLEVLDLNLNNTETVRNTYFPHGPPKELNLAGNRIGFLELGAFDGSLRSLLT 199
Qy 129 LEFDNRLLTTPNGAFVYLSKLKELWLRNNPIESPYAFNRIPSLRLDLGELKRLSYI 188
Db 200 LRLSKNRITQLPVRAF-KLPRLTOLDLNRNIRLIEGLTFQGLNSLVKL-QRNNISKL 257
Qy 189 SEGAFGLSNRLYLNAMCNLEI----- 212
Db 258 TDGAPWGLSKMHVHLHLEVDLSLEVNSGLYGLTALHQLHLSNNSIARIHRKGWSPCQKLH 317
Qy 213 -----PNLTPLIKLDELDSLGNHLSAIRPGSFQGLMHLO----- 246
Db 318 ELVLSFNLLRLDEESLAESSLVLRUSHNSISHIAGAFKGLSLRLVDLDHNEISGT 377
Qy 247 -----KLWIIQSQIQTIERNAFDNLQSLVEINLAHNNLTLLPHDLFTPL 290
Db 378 IEDTSGAFSGLEFGHSLKTLFGNKTLSVAKRAFSGLEGLHNLGNGNAIRSVOFDFVKM 437
Qy 291 HHLERIHLHNPWNCNDIILWSWI--KMAPSNATACARNTPPNLKGRYIGELDONQ 348
Db 438 KNLKELHISDSFLCDQCKWLPPMLGRMLQAFVATCAH---PESLKGQSIFSVPPES 494
Qy 349 FTC---YAPVIVEPPADLVNTEGMAELKCRASSTLS--VSWITPNGTVMTGCAKVR 402
Db 495 FVCDFLPKPIITPETTMAMVGKDIRFTCSAAASSSPMTFAWKDN-EVLTNADMENF 553
Qy 403 IAV-LSDG-----TLNFTNVVTDGTGMYTCMVNSVGNL---TASATLVNTAA-TTTP 450
Db 554 VHVHADGEVMEYITILHLRQVTFGHEGRYOCVITNHFHGSTYSHKARLTVNVLPSFTKTP 613

Qy 451 PSYFSTVTVETMEPSQDEARTDNNVGPVVDWETTNTVTSLTQPSTRSTEKTFIP-- 508
Db 614 HD-----ITRTTVARLECAATGH---PNQIAMQKDGTDPPAARERR---MHVMPDD 662
Qy 509 -----VTDINSIGPIDE-VMKTTKIIICGFVAITLMAAVMLVIFYKMKRQHRRHAPT 563
Db 663 DVFFITDVKIDDAGVYSCTAQNSAGSISANATLTVLETPSLVPLEDR----- 710
Qy 564 RTVEIINVDEI-----TGOTP-----MESHLPAPAEHEHNLHNYNSYKSPFNHTTVN 612
Db 711 -----VSVGETVALOCKATGNPPRITWFKGDRPLSLTERHHL-----TPDNQLLVQ 759
Qy 613 TI 614
Db 760 NV 761
RESULT 4
US-09-191-647-2
Sequence 2, Application US/09191647
Patent No. 6046015
GENERAL INFORMATION:
APPLICANT: Goodman, Corey
APPLICANT: Kid, Thomas
APPLICANT: Brose, Katja
APPLICANT: Tessier-Lavigne, Marc
TITLE OF INVENTION: Modulating Robo: Ligand Interactions
FILE REFERENCE: B98-031-3
CURRENT APPLICATION NUMBER: US/09/191,647
CURRENT FILING DATE: 1998-11-13
EARLIER APPLICATION NUMBER: 60/065,544
EARLIER FILING DATE: 1997-11-14
EARLIER APPLICATION NUMBER: 60/081,057
EARLIER FILING DATE: 1998-04-07
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2
LENGTH: 1525
TYPE: PRT
ORGANISM: human
US-09-191-647-2

Query Match 11.88; Score 395.5; DB 3; Length 1525;
Best Local Similarity 25.18; Pred. No. 6.7e-25;
Matches 129; Conservative 51; Mismatches 171; Indels 163; Gaps 9;
Qy 30 LIALQL---LVVAGLVRAQTCPSCVCSNQFSKVICVRKNLREVDPGISTNRLNLHEN 86
Db 8 MLSLSGLVLAILNKKVAPQACPAQCSCG--STVDCHGLALRSVPRNIPRNTERLDLNGN 65
Qy 87 QTOIIVNSFKHLRHLLETLQLSRNHIETIGAFNGLANLNTLELFDNRLLTTPNGAFVY 146
Db 66 NITRTKTDFAGLRHLVLQLMENKISTIERGAFODLELRLLNRNHLQFLPELLFIG 125
Qy 147 LSKLKLWLRNNPIESPYAFNRIPSLRLDLGELKRLSYISEGAFGLSNRLYNL-- 204
Db 126 TAKLYRLDLSENOIOAIRKAFRGAVDIKNLQL-DYNOISCIEDGAFRALRDLEVLTLNN 184
Qy 205 ----- 204
Db 185 NNITLSVASFNHMPKLRTRFLHSNNLYCDCHLAWLSDWLKRKRPRVGLYTQCMGFSHLRG 244
Qy 205 -----AMC-----NLREIN----- 214
Db 245 HNAEVQKREFVCSDEECHOFSMAPSCSVLHCPAACSNNIVDCRGKGLTEIPTNLPE 304
Qy 215 -----LTPLIKLDELDSLGNHLSAIRPGSFQ----- 240
Db 305 TITEIRLEQNTIKVIPPGAFSPYKKLRRIDLSNNQISELAPDAFQGLRSLNSLVYGNKI 364
Qy 241 -----GLMHLOKLMWIOSQIOVIERNAPDNLOSLVEINLAHNNLTLLPHDLFTPLH 291

Db 365 TELPKSLEGLFSLQLLLNANKINCLRVDAFQDLHNLNLLSLYDNKLTQIAKGTFSPLR 424
Qy 292 HLERIHLHNPNWNCNDILWLSWIKDMAPSNNTACCARCNTPPNKLGRYIGELDONYETC 351
Db 425 AIQTHLAQNPICDCHLKLWADYL-HTNPIETS-GARCTSPRLANKRIGOIKSKKPRC 482
Qy 352 YAPVIVEPPADLNVTEGMAELKCRASSTLSVS 385
Db 483 SGTEDYRSKLSGDCDFADLACPEKCRCEGTTVDCS 516
RESULT 5
US-09-540-245A-2
; Sequence 2, Application US/09540245A
; Patent No. 6270984
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey
; APPLICANT: Kid, Thomas
; APPLICANT: Brose, Katja
; APPLICANT: Tessier-Lavigne, Marc
; TITLE OF INVENTION: Modulating Robo: Ligand Interactions
; FILE REFERENCE: B98-031-3
; CURRENT APPLICATION NUMBER: US/09/540,245A
; CURRENT FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 60/065,544
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: 60/081,057
; PRIOR FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1525
; TYPE: PRT
; ORGANISM: human
US-09-540-245A-2
Query Match 11.8%; Score 395.5; DB 4; Length 1525;
Best Local Similarity 25.1%; Pred. No. 6.7e-25;
Matches 129; Conservative 51; Mismatches 171; Indels 163; Gaps 9;
Qy 30 LLALQL---LVVAGLVRAQTCPSCVSCSNQPSKVCVRKNLREVDPDGTSTNTRLNLHEN 86
Db 8 MLSLSGLVLVLAINKVAPOACPQCSG--STVDCHGLAURSVPRNIPRTERLDLNGN 65
Qy 87 QTIQIKVANSFKHLHLETLQLSRNHIRTIEGAFNGLANLNTLELFDNRLTTPNGAFVY 146
Db 66 NITRTKTDFAGLRHLRVLQLMENKISTIERGAFODLKELERLNRNHLQLFPPELLFLG 125
Qy 147 LSKKELMWRNPNIBSIPYAFNRIPSRRLDLGELKRLSYISEGAFGLSNRLYNL-- 204
Db 126 TAKLYRLDLSNQIAIPKAPRGAVDIKNLQL-DYNQISCTIEDGAFRALRDLEVLTLNN 184
Qy 205 ----- 204
Db 185 NNITRLSVASFNHMPKLTFRHLNNSNLYCDCHLAWSDWLKRPRVGLYTQCMGFSHLRG 244
Qy 205 ----- 214
Db 245 HNVAEQKREFVCSDEEGHOSFMAPSCSVLHCPAACTCSNNIVDCRGKGLTEIPTNLPE 304
Qy 215 ----- 240
Db 305 TITEIRLEQNTIKVPPGAFSPYKKLRRLDLSNNOISELAPDAFOGLRSLNSLVLYGNKI 364
Qy 241 ----- 291
Db 365 TELPKSLEGLFSLQLLLNANKINCLRVDAFQDLHNLNLLSLYDNKLTQIAKGTFSPLR 424
Qy 292 HLERIHLHNPNWNCNDILWLSWIKDMAPSNNTACCARCNTPPNKLGRYIGELDONYETC 351
Db 425 AIQTHLAQNPICDCHLKLWADYL-HTNPIETS-GARCTSPRLANKRIGOIKSKKPRC 482

Qy 352 YAPVIVEPPADLNVTEGMAELKCRASSTLSVS 385
Db 483 SGTEDYRSKLSGDCDFADLACPEKCRCEGTTVDCS 516
RESULT 6
US-09-540-153-2
; Sequence 2, Application US/09540153
; Patent No. 6270995
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey
; APPLICANT: Kid, Thomas
; APPLICANT: Brose, Katja
; APPLICANT: Tessier-Lavigne, Marc
; TITLE OF INVENTION: Modulating Robo: Ligand Interactions
; FILE REFERENCE: B98-031-3
; CURRENT APPLICATION NUMBER: US/09/540,153
; CURRENT FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/191,647
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: 60/081,057
; PRIOR FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1525
; TYPE: PRT
; ORGANISM: human
US-09-540-153-2
Query Match 11.8%; Score 395.5; DB 4; Length 1525;
Best Local Similarity 25.1%; Pred. No. 6.7e-25;
Matches 129; Conservative 51; Mismatches 171; Indels 163; Gaps 9;
Qy 30 LLALQL---LVVAGLVRAQTCPSCVSCSNQPSKVCVRKNLREVDPDGTSTNTRLNLHEN 86
Db 8 MLSLSGLVLVLAINKVAPOACPQCSG--STVDCHGLAURSVPRNIPRTERLDLNGN 65
Qy 87 QTIQIKVANSFKHLHLETLQLSRNHIRTIEGAFNGLANLNTLELFDNRLTTPNGAFVY 146
Db 66 NITRTKTDFAGLRHLRVLQLMENKISTIERGAFODLKELERLNRNHLQLFPPELLFLG 125
Qy 147 LSKKELMWRNPNIBSIPYAFNRIPSRRLDLGELKRLSYISEGAFGLSNRLYNL-- 204
Db 126 TAKLYRLDLSNQIAIPKAPRGAVDIKNLQL-DYNQISCTIEDGAFRALRDLEVLTLNN 184
Qy 205 ----- 204
Db 185 NNITRLSVASFNHMPKLTFRHLNNSNLYCDCHLAWSDWLKRPRVGLYTQCMGFSHLRG 244
Qy 205 ----- 214
Db 245 HNVAEQKREFVCSDEEGHOSFMAPSCSVLHCPAACTCSNNIVDCRGKGLTEIPTNLPE 304
Qy 215 ----- 240
Db 305 TITEIRLEQNTIKVPPGAFSPYKKLRRLDLSNNOISELAPDAFOGLRSLNSLVLYGNKI 364
Qy 241 ----- 291
Db 365 TELPKSLEGLFSLQLLLNANKINCLRVDAFQDLHNLNLLSLYDNKLTQIAKGTFSPLR 424
Qy 292 HLERIHLHNPNWNCNDILWLSWIKDMAPSNNTACCARCNTPPNKLGRYIGELDONYETC 351
Db 425 AIQTHLAQNPICDCHLKLWADYL-HTNPIETS-GARCTSPRLANKRIGOIKSKKPRC 482
Qy 352 YAPVIVEPPADLNVTEGMAELKCRASSTLSVS 385
Db 483 SGTEDYRSKLSGDCDFADLACPEKCRCEGTTVDCS 516
RESULT 7

QY 46 TCPSVCSNQSFCVCRKRLNREVPDGIISNTNRLNLHENQIQIKVNSFKHLRHLEIL 105
Db 294 SCPPHRCADGI--VDCREKSLTSPVTPDPTDTRVLEQNFITELPKFSFRRRLRI 351
QY 106 QLSRNRHRTTEIGAFNGLANLNTLELFDNRLLTTPNGAFVYLSKLKELWLRNPNIESIPS 165
Db 352 DLSNNISRTAHDALSGLKQLTTLVLVGNKIKDLPVGFGLGSLRLRLNANEISCIK 411
QY 166 YAFNRIPSLRRLDGLKRLSYISEGAFEGLSNRLYNLAW-----CNLR----- 210
Db 412 DAFRDLHSLLSLYD--NNIQSLANGTFDAMKSMKTVHLAKNPPFICDNCNRLWLADYLHN 470
QY 211 -----EIPNLTPPL-----IKLD-----ELDLSGNHLSAIRPGS 238
Db 471 PIETSGARCESPKRMHRRRIETSLREKFKCSWGLRMKLSGECRMDSDCPAMCHCEGTV 530
QY 211 -----EIPNLTPPL-----IKLD-----ELDLSGNHLSAIRPGS 238
Db 531 DCTGRRRLKEPRDIPHTTELLNDNLGRISSDGLFGRPLHLVKLEKRNQLTGIEPNA 590
QY 239 FQGLMHLOKLMWIOSQIOVIERNAPDNLOSIVEINLAHNNTLLPLHDLFTPLHHLRIHL 298
Db 591 FEGASHIOELQGLGENKIKESNKMFLGLHQLKTLNLYDNQISCVMPGSFEHLNLSLNL 650
QY 299 HHNPWNCNDILWLSWIKDWPNTACCARCNPPLNKGRIYIGELDONFTCYAPVIVE 358
Db 651 ASNPFNCNCHLAWFAECVRKSLNGGA--ARCGAPSKVRDQVQIKDLPHSEFKCSS----- 703
QY 359 PPADLVNTEGMAAELKCRASLTSLTSVSWITPNGTVM 394
Db 704 -----ENSEGCLGDGYPSCCTCT-----GTVV 726

RESULT 12

US-09-540-153-7
; Sequence 7, Application US/09540153
; Patent No. 6270995
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey
; APPLICANT: Kid, Thomas
; APPLICANT: Brose, Katja
; APPLICANT: Tessier-Lavigne, Marc
; TITLE OF INVENTION: Modulating Robo: Ligand Interactions
; FILE REFERENCE: B98-031-3
; CURRENT APPLICATION NUMBER: US/09/540,153
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/191,647
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: 60/081,057
; PRIOR FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1480
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-540-153-7

Query Match 10.4%; Score 350; DB 4; Length 1480;
Best Local Similarity 23.9%; Pred. No. 5.1e-21;
Matches 109; Conservative 65; Mismatches 152; Indels 130; Gaps 9;
QY 46 TCPSVCSNQSFCVCRKRLNREVPDGIISNTNRLNLHENQIQIKVNSFKHLRHLEIL 105
Db 294 SCPPHRCADGI--VDCREKSLTSPVTPDPTDTRVLEQNFITELPKFSFRRRLRI 351
QY 106 QLSRNRHRTTEIGAFNGLANLNTLELFDNRLLTTPNGAFVYLSKLKELWLRNPNIESIPS 165
Db 352 DLSNNISRTAHDALSGLKQLTTLVLVGNKIKDLPVGFGLGSLRLRLNANEISCIK 411
QY 166 YAFNRIPSLRRLDGLKRLSYISEGAFEGLSNRLYNLAW-----CNLR----- 210

Db 412 DAFRDLHSLLSLYD--NNIQSLANGTFDAMKSMKTVHLAKNPPFICDNCNRLWLADYLHN 470
QY 211 -----EIPNLTPPL-----IKLD-----ELDLSGNHLSAIRPGS 238
Db 471 PIETSGARCESPKRMHRRRIETSLREKFKCSWGLRMKLSGECRMDSDCPAMCHCEGTV 530
QY 211 -----EIPNLTPPL-----IKLD-----ELDLSGNHLSAIRPGS 238
Db 531 DCTGRRRLKEPRDIPHTTELLNDNLGRISSDGLFGRPLHLVKLEKRNQLTGIEPNA 590
QY 239 FQGLMHLOKLMWIOSQIOVIERNAPDNLOSIVEINLAHNNTLLPLHDLFTPLHHLRIHL 298
Db 591 FEGASHIOELQGLGENKIKESNKMFLGLHQLKTLNLYDNQISCVMPGSFEHLNLSLNL 650
QY 299 HHNPWNCNDILWLSWIKDWPNTACCARCNPPLNKGRIYIGELDONFTCYAPVIVE 358
Db 651 ASNPFNCNCHLAWFAECVRKSLNGGA--ARCGAPSKVRDQVQIKDLPHSEFKCSS----- 703
QY 359 PPADLVNTEGMAAELKCRASLTSLTSVSWITPNGTVM 394
Db 704 -----ENSEGCLGDGYPSCCTCT-----GTVV 726

RESULT 13

PCT-US91-09055-2
; Sequence 2, Application PC/TUS9109055
; GENERAL INFORMATION:
; APPLICANT: Rothberg, Jonathan Marc and Artavanis-Tsakonas, Spyridon
; TITLE OF INVENTION: Purified SLIT protein and Sequence Elements Thereof
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Yale University
; ADDRESSEE: Office of Cooperative Research
; STREET: 246 Church Street
; STREET: Suite 401
; CITY: New Haven,
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06510
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 800 Kb storage
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 6.0.5
; SOFTWARE: Microsoft Word 4.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/09055
; FILING DATE: 19911127
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/624,135
; FILING DATE: 7-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Barth, Richard J.
; REGISTRATION NUMBER: 28,180
; REFERENCE/DOCKET NUMBER: 900964/RSB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 972-1400
; TELEFAX: (212) 370-1622
; TELEX: 236268
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1480 amino acids
; TYPE: AMINO ACIDS
; TOPOLOGY: Linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: signal sequence
; LOCATION: 1 to 36
; IDENTIFICATION METHOD: similarity to other signal
; OTHER INFORMATION: Directs Export
; NAME/KEY: Four Flank-LRR-Flank domains
; LOCATION: 37 to 910
; IDENTIFICATION METHOD: Array of Flank-LRR-Flank

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 15:41:53 ; Search time 31.57 Seconds

(without alignments)

1681.764 Million cell updates/sec

Title: US-09-905-056-292_COPY_45_522

Perfect score: 2524

Sequence: 1 QTCPSVCSNQFSKVICVR.....KTFTPTVDINSIGIPIDEV 478

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*

```
1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2524	100.0	640	20	Human PRO331 prote
2	2524	100.0	640	20	Novel protein (Clo
3	2524	100.0	640	20	Amino acid sequenc
4	2524	100.0	640	21	Human PRO331 prote
5	2524	100.0	640	21	Human PRO331 prote
6	2524	100.0	640	22	Human immune respo
7	2524	100.0	640	22	Human PRO331 prote
8	2524	100.0	640	22	Human PRO331 prote
9	2524	100.0	640	22	Human PRO331 prote
10	2524	100.0	640	22	Human angiogenesis
11	1793.5	71.1	713	22	Human leucine-rich

12	1581.5	62.7	653	20	AA28806	cc359_4 secreted p
13	1581.5	62.7	653	21	AA24073	Human PRO1111 prot
14	1581.5	62.7	653	21	AA66694	Membrane-bound pro
15	1581.5	62.7	653	22	AA09438	Human sbgPRO331a p
16	1581.5	62.7	653	22	AAU12390	Human PRO1111 poly
17	1581.5	62.7	653	22	AA65217	Human PRO1111 (UNQ
18	1561.5	61.9	448	21	AA23044	Human SLIT protein
19	1560.5	61.8	406	22	AAU18035	Human immunoglobul
20	1548.5	61.4	590	21	AA23034	Human SLIT protein
21	1548.5	61.4	694	21	AA23033	Human SLIT protein
22	1356.5	53.7	553	22	AB10349	Human CDNA SEQ ID
23	971	38.5	302	21	AA87066	Human secreted pro
24	971	38.5	302	22	AAE06043	Human gene 3 encod
25	886.5	35.1	441	21	AA43091	Human ORFX ORF2855
26	845.5	33.5	281	22	AB10522	Human CDNA SEQ ID
27	845.5	33.5	281	22	AAU18084	Human immunoglobul
28	685	27.1	224	21	AA87135	Human secreted pro
29	685	27.1	224	22	AAE06112	Human gene 3 encod
30	568.5	22.5	614	20	AAW84596	Amino acid sequenc
31	565.5	22.4	579	22	AAE06804	Mature human neuro
32	565.5	22.4	606	22	AAE06799	Human neuronal gui
33	565.5	22.4	620	22	AAE74705	Human membrane ass
34	556.5	22.0	620	20	AA13357	Amino acid sequenc
35	556.5	22.0	620	22	AAU12333	Human PRO227 poly
36	556.5	22.0	620	22	AAE00225	Human PRO227 prote
37	522.5	20.7	592	22	AAE09437	Human sbgTango79a
38	477.5	18.9	548	22	AAE31161	Amino acid sequenc
39	453	17.9	766	22	AAE67505	Amino acid sequenc
40	443.5	17.6	1021	22	AB860162	Drosophila melanog
41	430.5	17.1	713	20	AA13385	Amino acid sequenc
42	430.5	17.1	713	22	AAE00253	Human PRO293 prote
43	422.5	16.7	707	20	AAU02379	Polypeptide ident
44	421	16.7	705	21	AAE42513	Human ORFX ORF2277
45	421	16.7	705	22	AAW78823	Human protein SEQ

ALIGNMENTS

```
RESULT 1
AA08100
ID AA08100 standard; Protein; 640 AA.
XX
AC AA08100;
XX
DT 11-SEP-2000 (first entry)
XX
DE Human PRO331 protein.
XX
```

Inflammatory cell infiltration; immune response; T cell proliferation;
anti-inflammatory; anti-autoimmune; anti-diabetic; spondyloarthritis;
T cell-mediated disease; spondyloarthritis; sclerosis; renal disease;
inflammatory myopathy; hemolytic anemia; thrombocytopenia; thyroiditis;
diabetes mellitus; demyelinating polyneuropathy; Gullain-Barre syndrome;
multiple sclerosis; polyneuropathy; hepatitis; cirrhosis; enteropathy;
sclerosing cholangitis; inflammatory bowel disease; Whipple's disease;
skin disease; dermatitis; psoriasis; asthma; allergic rhinitis; tumor;
food hypersensitivity; urticaria; eosinophilic pneumonia; transplant;
idiopathic pulmonary fibrosis; graft rejection; PRO245; human;
PRO331.

```
XX Homo sapiens.
XX
XX WO9914241-A2.
XX
XX 25-MAR-1999.
XX
XX 17-SEP-1998; 98WO-US19437.
XX
XX 17-SEP-1997; 97US-0059119.
XX 18-SEP-1997; 97US-0059263.
XX 28-OCT-1997; 97US-0063550.
XX 12-NOV-1997; 97US-0065186.
```

PR 21-NOV-1997; 97US-0066364.
PR 24-NOV-1997; 97US-0066770.
PR 04-JUN-1998; 98US-0088026.
PA (GETH) GENENTECH INC.
XX
PI Fong S, Goddard A, Gurney AL, Tumas D, Wood WI;
XX WPI; 1999-229499/19.
DR N-PSDB; AAX37738.
XX
XX Composition containing novel polypeptide PRO245, its agonist or
PT antagonist -
XX
XX Example 1; Fig 33; 177pp; English.
XX
XX This invention describes a novel composition containing (apart from a
CC carrier or excipient), a novel PRO245 polypeptide (I), its agonist or
CC antagonist, or their fragments, for modulating: (i) infiltration of
CC inflammatory cells into tissue; (ii) an immune response; or (iii) T cell
CC proliferation. The composition increases or decreases any of the effects
CC (i)-(iii). The products of the invention have anti-inflammatory,
CC anti-autoimmune and anti-diabetic activity. (I), and its (ant)agonists
CC and their fragments, are used to treat immune-related diseases,
CC particularly T cell-mediated diseases. The diseases treated include
CC systemic lupus erythematosus, rheumatoid arthritis, juvenile chronic
CC arthritis, spondyloarthropathies, systemic sclerosis (scleroderma),
CC idiopathic inflammatory myopathies (dermatomyositis, polymyositis),
CC Sjogren's syndrome, systemic vasculitis (dermatomyositis, polymyositis),
CC hemolytic anemia (immune pancytopenia, paroxysmal nocturnal
CC hemoglobinuria), autoimmune thrombocytopenia (idiopathic thrombocytopenic
CC purpura), immune-mediated thrombocytopenia), thyroiditis (Grave's disease,
CC Hashimoto's thyroiditis, juvenile lymphocytic thyroiditis, atrophic
CC thyroiditis), diabetes mellitus, immune-mediated renal disease
CC (glomerulonephritis, tubulointerstitial nephritis), multiple sclerosis,
CC idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome, chronic
CC inflammatory demyelinating polyneuropathy, infectious hepatitis
CC (hepatitis A, B, C, D, E and other non-hepatotropic viruses), autoimmune
CC chronic active hepatitis, primary biliary cirrhosis, granulomatous
CC hepatitis, and sclerosing cholangitis, inflammatory bowel disease
CC (ulcerative colitis; Crohn's disease), gluten-sensitive enteropathy, and
CC Whipple's disease. Autoimmune or immune-mediated skin diseases including
CC bullous skin diseases, erythema multiforme, contact dermatitis, psoriasis,
CC asthma, allergic rhinitis, atopic dermatitis, food hypersensitivity,
CC urticaria, eosinophilic pneumonia, idiopathic pulmonary fibrosis,
CC hypersensitivity pneumonitis, and transplantation associated diseases
CC (graft rejection, and graft-versus-host-disease). (I), its (ant)agonists
CC or fragment can also be used as an adjuvant in treatment of tumors.
CC Antibodies against (I) can also be used for diagnosing such diseases.
CC This sequence represents the human PRO331 protein which is described in
CC the method of the invention.
XX
SQ Sequence 640 AA;

Query Match 100.0%; Score 2524; DB 20; Length 640;
Best Local Similarity 100.0%; Pred. No. 2.4e-177;
Matches 478; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QTCPSVCSNSQFSKVCVRKRLREVPDGIQSTNRLNLNHNQIITKIVNSFKHLRLEI 60
Db 45 qtcpsvcsnsqfskvcvrkrlrevpdgiqstnrlnlhnenqiklqvnsfkhlrlel 104
QY 61 LQLSRNHRTIEGAFNGLANLTLEFDNRLTIPNGAFVYLSKLKELWLRNPPIESIP 120
Db 105 lqlsrnhrtiegaflnglanltlefdnrlltippngafvylskelwlrnpiesip 164
QY 121 SYANRNPISRLRDLGELKRLSYISEGAFEGLSNRLYNLNCMLREIPNTPILKIDEL 180
Db 165 syanrnpisrlrdlgelkrlsyisegafeglsnrlynlncmlreipnltplkidel 224
QY 181 DLSGNHLSAIRPGSFGQHLQKLMWQSQIQVTERNAFNLQSLVEINLAHNNLTLLPH 240
XX

Db 225 dlsghnhsairpgsfqglmhlikqlwmiqsgiqviernafndnlgsiveinlahnnltllph 284
QY 241 DLFTPLHLHLRIHLHHPNWCNCDILWLSWIKDMAPSNNTACCARCNPPLNKGRIYIGEL 300
Db 285 dlftplhlhlrlhlhnpwnncndilwlswwikdmapsntaccarcncppnlkgryigel 344
QY 301 DQNYFTCYAPVIVPEPPADLNVTGMAELKCRASSTLSVSWITPNGVTMTHGAYKVRIA 360
Db 345 dqnyftcyapviveppadlnvtgmaaelkcrastlsvswitpngvtmthgaykvria 404
QY 361 VLSGTLNFTNVTVDGMYTCMVNSVGNVTASATLNVTAAATTPPFSYFSTVVTVEP 420
Db 405 vlsdgtlnftnvtvgdgmtytcmvnsvgnvtasatlnvtaatttppfsyfstvtvetnep 464
QY 421 SODEARTDNNVGPVPVVDWETTNVTTSLTPQSTRSTKTTIPVTDNSGIPGIDEV 478
Db 465 sqdeartdnnvgpvpvvdwettnttsltbpqstrstekttipvtdnsgipgidev 522
RESULT 2
AAM85722
ID AAM85722 standard; Protein; 640 AA.
XX
AC AAM85722;
XX
XX 27-SEP-1999 (first entry)
XX
XX Novel protein (Clone AS209_1).
XX
KW Polynucleotide; protein; nutrition; cytokine; cell proliferation;
KW cell differentiation; immunostimulation; immunosuppression;
KW hematopoiesis regulation; tissue growth; activin; inhibin;
KW chemotaxis; chemokinesis; haemostasis; thrombolysis; receptor;
KW ligand; anti-inflammatory; tumour suppression; gene therapy.
XX
OS Homo sapiens.
XX
XX WO9920644-A1.
XX
XX 29-APR-1999.
XX
XX 16-OCT-1998; 98WO-US22034.
XX
XX 18-OCT-1997; 97US-0955557.
XX
XX (GEMY) GENETICS INST INC.
XX
XX Agostino MJ, Bowman MR, Evans C, Jacobs K, Lavallie ER;
XX McCoy JM, Merberg D, Racie LA, Spaulding V, Treacy M;
XX WPI; 1999-288272/24.
XX
XX N-PSDB; AAX08687.
XX
XX New polynucleotides encoding secreted human proteins
PT
XX
XX Claim 26; Page 109-111; 136pp; English.
XX
XX The new human secreted proteins are encoded by polynucleotides
XX obtained from human placenta, adult testes, fetal kidney, fetal
XX brain, adult brain, adult brain and adult blood cDNA libraries.
XX The polynucleotides and proteins are predicted to have biological
XX activities which would make them suitable for treating, preventing or
XX ameliorating medical conditions in humans and animals. Suggested
XX activities include nutritional activity, cytokine and cell
XX proliferation/differentiation activity, immune stimulating (e.g. as
XX vaccines) or suppressing activity, haematopoiesis regulating
XX activity, tissue growth activity, activin/inhibin activity,
XX chemotactic/chemokinetic activity, haemostatic and thrombolytic
XX activity, receptor/ligand activity, anti-inflammatory activity,
XX cadherin/tumour invasion suppressor activity, and tumour inhibition
XX activity. The polynucleotides are also stated to be useful for gene
XX therapy. The sequences identified by a secretory leader
XX sequence motif in the polynucleotide and it is thought that the

CC encoded proteins have biological activity by virtue of their secreted nature. This polypeptide was encoded by a clone designated A3209_1 (See AAX08687).

XX Sequence 640 AA;

Query Match 100.0%; Score 2524; DB 20; Length 640;
Best Local Similarity 100.0%; Pred. No. 2.4e-177;
Matches 478; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QTCPSCSCNQSKVICVRKNLREVPDGIISTNRLNLHENOIQIKVNSFKHLRHL 60
DB 45 qtcpsscscnqskvicvrknirevpdgistnrlnlhendqikvnsfkhlrhlei 104
QY 61 LQLSNRHIRTIEGAFNGLANLTLEFDNRLTTPNGAFVYLSKELWLRNPNIESIP 120
DB 105 lqlsnrhirtieigafnglanltlefdnrltppngafvyskikewlrnnpiesip 164
QY 121 SYAFNRIPSLRRDLGELKRLSYISEGAFGLSNLRYLNAMCNLRPIPLIKLDEL 180
DB 165 syafnrslrrldgelkrlsyisegafeglsnlylnlamcnlreipnltplikidel 224
QY 181 DLGSHLSAIRPGSFQGLMHLQKLWIKQSIQVIERNAFNLQSLVEINLAHNLLPLH 240
DB 225 dlsgnhsairpgsfqglmhlqklwmiqsiqviernafndlqslveinlahnlltlph 284
QY 241 DLFTPLHLERILHHPWNCNDILWLSWIKDMAPSNACCARCNTPPNLKGRYIGEL 300
DB 285 dlftplhlrlhlpwncndilwlswwikdmapsnaccarcntppnlkgrylgel 344
QY 301 DQNYFTCYAPVIEPPADLNVTEGMAELKCRASSTLSVSWITPNTGVTMTHGAYKVRIA 360
DB 345 dqnyftcyapvieveppadlnvtegmaaelkcrasstlsvswitpntgvtmthgaykria 404
QY 361 VLSGTLNFTNVVQDTGMYTCMVNSVGNVTASALNVTAATTPSPSYSTVVTMEP 420
DB 405 vlsdgtlnftnvtvqdtgmytcmvnsvgnttasatlnvtaattppsfystvtvmetep 464
QY 421 SQDEARTDNNVGPVVDWETTNVTLSLPQSTRSEKFTTIPVTDINSIGIPIDEV 478
DB 465 sqdeartdnnvgpvpvdwettntvtsltpqstrsektftipvtidinsigipidev 522

RESULT 3

AAY13394
ID AAY13394 standard; Protein; 640 AA.

XX AAY13394;

XX 25-JUN-1999 (first entry)

XX Amino acid sequence of protein PRO331.

XX Secreted protein; transmembrane protein; human; enterocolitis;
KW Zollinger-Ellison syndrome; gastrointestinal ulceration;
KW congenital microvillus atrophy; skin disease; cell growth;
KW abnormal keratinocyte differentiation; psoriasis; epithelial cancer;
KW Parkinson's disease; Alzheimer's disease; ALS; neuropathy;
KW fibromodulin; dermal scarring; Usher Syndrome; Atrophia areata;
KW anti-thrombotic; wound healing; tissue repair.

XX Homo sapiens.

XX WO9914328-A2.

XX 25-MAR-1999.

XX 16-SEP-1998; 98WO-US19330.

XX 25-NOV-1997; 97US-0066840.

XX 17-SEP-1997; 97US-0059113.

XX 17-SEP-1997; 97US-0059115.

PR 17-SEP-1997; 97US-0059117.
PR 17-SEP-1997; 97US-0059119.
PR 17-SEP-1997; 97US-0059121.
PR 17-SEP-1997; 97US-0059122.
PR 17-SEP-1997; 97US-0059184.
PR 18-SEP-1997; 97US-0059263.
PR 18-SEP-1997; 97US-0059266.
PR 15-OCT-1997; 97US-0062125.
PR 17-OCT-1997; 97US-0062285.
PR 17-OCT-1997; 97US-0062287.
PR 21-OCT-1997; 97US-0063486.
PR 24-OCT-1997; 97US-0062814.
PR 24-OCT-1997; 97US-0062816.
PR 24-OCT-1997; 97US-0063045.
PR 24-OCT-1997; 97US-0063120.
PR 24-OCT-1997; 97US-0063121.
PR 24-OCT-1997; 97US-0063127.
PR 24-OCT-1997; 97US-0063128.
PR 27-OCT-1997; 97US-0063329.
PR 27-OCT-1997; 97US-0063327.
PR 28-OCT-1997; 97US-0063541.
PR 28-OCT-1997; 97US-0063542.
PR 28-OCT-1997; 97US-0063544.
PR 28-OCT-1997; 97US-0063549.
PR 28-OCT-1997; 97US-0063550.
PR 28-OCT-1997; 97US-0063564.
PR 29-OCT-1997; 97US-0063435.
PR 29-OCT-1997; 97US-0063704.
PR 29-OCT-1997; 97US-0063732.
PR 29-OCT-1997; 97US-0063738.
PR 29-OCT-1997; 97US-0063734.
PR 29-OCT-1997; 97US-0064215.
PR 31-OCT-1997; 97US-0063735.
PR 31-OCT-1997; 97US-0063870.
PR 31-OCT-1997; 97US-0064103.
PR 03-NOV-1997; 97US-0064248.
PR 07-NOV-1997; 97US-0064809.
PR 12-NOV-1997; 97US-0065186.
PR 17-NOV-1997; 97US-0065846.
PR 18-NOV-1997; 97US-0065693.
PR 21-NOV-1997; 97US-0066120.
PR 21-NOV-1997; 97US-0066364.
PR 24-NOV-1997; 97US-0066772.
PR 24-NOV-1997; 97US-0066466.
PR 24-NOV-1997; 97US-0066770.
PR 24-NOV-1997; 97US-0066511.
PR 24-NOV-1997; 97US-0066453.

(GETH) GENENTECH INC.

Chen J, Goddard A, Gurney AL, Pennica D, Wood WI, Yuan J;

WPI; 1999-229533/19.

N-PSDB; AAX52265.

New isolated human genes and polypeptides used in, e.g. treatment of gastrointestinal ulceration

Claim 12; Fig 104; 320pp; English.

AA13344-403 represent secreted and transmembrane human proteins.

The cDNA sequences are obtained from cDNA libraries, prepared from fetal lung, fetal kidney, fetal brain, fetal liver and fetal retina.

The encoded polypeptides have specific uses based on their homology to known polypeptides, e.g. PRO211 and PRO217 can be used for disorders associated with the preservation and maintenance of gastrointestinal mucosa and the repair of acute and chronic mucosal lesions (e.g. enterocolitis, Zollinger-Ellison syndrome, gastrintestinal ulceration and congenital microvillus atrophy), skin diseases associated with abnormal keratinocyte differentiation (e.g. psoriasis, epithelial cancers such as lung squamous cell carcinoma of the vulva and gliomas), potent effects on cell growth and development, diseases related to growth or survival of nerve cells including Parkinson's disease,

CC Alzheimer's disease, ALS, neuropathies or cancer. PRO265 can be used as
CC for fibromodulin, e.g. for reducing dermal scarring. PRO264 can be used
CC as a target for anti-tumor drugs. PRO533 may be used in the treatment
CC of Usher Syndrome or Atrophia areata; PRO269 can be used as an
CC anti-thrombotic agent; PRO287 polypeptides and portions may have
CC therapeutic applications in wound healing and tissue repair; PRO317 can
CC be used for treating problems of the kidney, uterus, endometrium, blood
CC vessels, or related tissue, e.g. in the heart of genital tract.

XX Sequence 640 AA;

Query Match 100.0%; Score 2524; DB 20; Length 640;
Best Local Similarity 100.0%; Pred. No. 2.4e-177;
Matches 478; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QTCPSVCSNQSFKVTCVRKNLREVPDGI STNTRLLNHNHQI I K V N S F K H L R H L E I 60
Db 45 qtcpvscsnqfskvcvcrknrlrevpdg istntrllnhenq i k l v n s f k h l r h l e i 104
QY 61 LQSRNHRTIEIGAFNGLANLTLELFDNRLTTPNGAFVYLSKELWLRNPIESIP 120
Db 105 lqsrnhrtieigafnglanltlelfdnrlttipngafv y l s k e l w l r n p i e s i p 164
QY 121 SYAFNRIPSLRRDLGELKRLSYISEGAFGLSNLRYLNAMCNLREIPNLTPLIKDEL 180
Db 165 syafnrpslrrldgelkrlsyisegafeglsn l r y l n a m c n l r e i p n l t p l i k d e l 224
QY 181 DLGNHLSAIRPGSFQGLMHQLKLMWIOSQIOVIERNAFDNLQSLVEINLAHNLTLLPH 240
Db 225 dlsghlsairpgsfqglmhqlklmwiqslqviernaf d n l q s l v e i n l a h n l t l l p h 284
QY 241 DLFTPLHLERIHLPNPNWNCNCDILWLSWIKDMAPSNTACCRCNTPPNKLGRYIGEL 300
Db 285 dlftplhlhrlhlpnpnwnncd ilwlswwikdmapsn t a c c r c n t p p n k l g r y i g e l 344
QY 301 DQNTFYCAPVIVPEPPADLNVTGMAAELKCRASTSLTSVSWITPNGTVMTGHGAYKRIA 360
Db 345 dqnyftcyapviveppadlnvtgmaaelkcrastsl t s v s w i t p n g t v m t g h g a y k r i a 404
QY 361 VLSQGTINFTNVTQDGMVTCMVSNVSGNTASATLNVTAAATTPFSYFSTVETMEP 420
Db 405 vlsqgtinfntvtdqgmvcmsvnsngntasatlnv t a a t t p f s y f s t v e t m e p 464
QY 421 SQDEARTDNNVGPVVDVETTNVTSLTPQSTRSTEKFTTIPVTDINSIGIPGIDEV 478
Db 465 sqdeartdnnvgptvvdvettntvtstltpqstrst e k f t t i p v t d i n s i g i p g i d e v 522

RESULT 4

AAB24407
ID AAB24407 standard; Protein; 640 AA.
XX
AC AAB24407;
XX
DT 07-NOV-2000 (first entry)
DE Human PRO331 protein sequence SEQ ID NO:107.
XX
KW Human; PRO; promotion; inhibition; angiogenesis; cardiovascularisation;
KW diagnosis; trauma; wound; cancer; atherosclerosis; cardiac hypertrophy;
KW angiogenic; proliferative; cardiant; cardiovascular; antiatherosclerotic;
KW cytostatic; gene therapy; vaccine.
XX
OS Homo sapiens.
XX
PN W0200032221-A2.
XX
PD 08-JUN-2000.
XX
PF 30-NOV-1999; 99WO-US28313.
XX
PR 01-DEC-1998; 98WO-US25108.

PR 16-DEC-1998; 98US-0112850.
PR 12-JAN-1999; 99US-0115554.
PR 08-MAR-1999; 99WO-US05028.
PR 12-MAR-1999; 99US-0123957.
PR 28-APR-1999; 99US-0131445.
PR 14-MAY-1999; 99US-0134287.
PR 02-JUN-1999; 99WO-US12252.
PR 23-JUN-1999; 99US-0141037.
PR 20-JUL-1999; 99US-0144758.
PR 26-JUL-1999; 99US-0145698.
PR 01-SEP-1999; 99WO-US20111.
PR 08-SEP-1999; 99WO-US20594.
PR 13-SEP-1999; 99WO-US20944.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 03-OCT-1999; 99WO-US23089.
PR 29-OCT-1999; 99US-0162506.
XX
PA (GETH) GENENTECH INC.

XX Ashkenazi AJ, Baker KP, Ferrara N, Gerber H, Hillan KJ, Goddard A;
PI Godowski PJ, Gurney AL, Klein RD, Kuo SS, Paoni NF, Smith V;
PI Watanabe CK, Williams PM, Wood WI;
XX
DR WPI; 2000-412154/35.
DR N-PSDB; AAA77596.

XX Nucleic acids encoding PRO polypeptides useful for preventing,
PT diagnosing and treating disorders a cardiovascular, endothelial or
PT angiogenic disorders in mammals -
XX
PS Claim 72; Fig 40; 315pp; English.

XX The present invention describes nucleic acids encoding PRO polypeptides
CC useful for preventing, diagnosing and treating disorders a
CC cardiovascular, endothelial or angiogenic disorder in mammals by
CC modulating cell proliferation, angiogenesis and cardiovascularisation,
CC and for identifying agonists and antagonists of these processes. The
CC nucleic acids and the proteins they encode may be used in the
CC prevention, treatment and diagnosis of diseases associated with
CC inappropriate PRO expression such as cardiovascular, endothelial or
CC angiogenic disorders in mammals (e.g. atherosclerosis, cancers and
CC cardiac hypertrophy). For example, the nucleic acids (NAs) and vectors
CC containing them and the PRO polypeptide may be used to treat disorders
CC associated with decreased PRO expression. AAA77510 to AAA77721 and
CC AAB24388 to AAB24435 represent nucleotide and protein sequences used in
CC the exemplification of the present invention.

XX Sequence 640 AA;

Query Match 100.0%; Score 2524; DB 21; Length 640;
Best Local Similarity 100.0%; Pred. No. 2.4e-177;
Matches 478; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QTCPSVCSNQSFKVTCVRKNLREVPDGI STNTRLLNHNHQI I K V N S F K H L R H L E I 60
Db 45 qtcpvscsnqfskvcvcrknrlrevpdg istntrllnhenq i k l v n s f k h l r h l e i 104
QY 61 LQSRNHRTIEIGAFNGLANLTLELFDNRLTTPNGAFVYLSKELWLRNPIESIP 120
Db 105 lqsrnhrtieigafnglanltlelfdnrlttipngafv y l s k e l w l r n p i e s i p 164
QY 121 SYAFNRIPSLRRDLGELKRLSYISEGAFGLSNLRYLNAMCNLREIPNLTPLIKDEL 180
Db 165 syafnrpslrrldgelkrlsyisegafeglsn l r y l n a m c n l r e i p n l t p l i k d e l 224
QY 181 DLGNHLSAIRPGSFQGLMHQLKLMWIOSQIOVIERNAFDNLQSLVEINLAHNLTLLPH 240
Db 225 dlsghlsairpgsfqglmhqlklmwiqslqviernaf d n l q s l v e i n l a h n l t l l p h 284
QY 241 DLFTPLHLERIHLPNPNWNCNCDILWLSWIKDMAPSNTACCRCNTPPNKLGRYIGEL 300
XX

Db 285 dlftplhlhlnhpncndllwskidmapsntaccarentppnlgkryigel 344
AY70673
ID AAY70673 standard; Protein; 640 AA.
AC AAY70673;
XX
DT 18-JUL-2000 (first entry)
XX
DE Human PRO331 protein.
XX
KW PRO331; UNQ292; dermatological; immunosuppressive; antiinflammatory;
immunostimulant; antilasthmatic; antirheumatic; antiarthritic; virucide;
antiallergic; haemostatic; hepatotropic; antidiabetic; antianaemic;
neurotropic; neuroprotective; anticoagulant; immunological disorder;
lung; pneumonia; skin; psoriasis; kidney; glomerulonephritis; arthritis;
spondyloarthropathy; SLE; systemic lupus erythematosus; scleroderma;
idiopathic inflammatory myopathy; anaemia; thrombocytopenia; diabetes;
thyroiditis; Grave's disease; demyelinating disease; multiple sclerosis;
Crohn's disease; hepatobiliary disease; hepatitis; asthma; human;
graft-versus-host-disease.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 40..46
FT /note= "N-myristoylation site"
FT Modified-site 73..79
FT /note= "N-myristoylation site"
FT Modified-site 118..124
FT /note= "N-myristoylation site"
FT Modified-site 183..187
FT /note= "cAMP and cGMP-dependent protein kinase
phosphorylation site"
FT Modified-site 191..197
FT /note= "N-myristoylation site"
FT Modified-site 228..234
FT /note= "N-myristoylation site"
FT Modified-site 237..243
FT /note= "N-myristoylation site"
FT Modified-site 268..272
FT /note= "Casein Kinase II phosphorylation site"
FT Modified-site 278..282
FT /note= "N-glycosylation site"
FT Modified-site 364..368
FT /note= "N-glycosylation site"
FT Modified-site 390..394
FT /note= "N-glycosylation site"
FT Modified-site 391..397
FT /note= "N-myristoylation site"
FT Modified-site 412..416
FT /note= "N-glycosylation site"
FT Modified-site 415..419
FT /note= "N-glycosylation site"
FT Modified-site 417..421
FT /note= "Casein Kinase II phosphorylation site"
FT Modified-site 422..428
FT /note= "N-myristoylation site"
FT Modified-site 433..439

FT Modified-site /note= "N-myristoylation site"
FT 434..438
FT /note= "N-glycosylation site"
FT 442..446
FT /note= "N-glycosylation site"
FT 465..469
FT /note= "Casein Kinase II phosphorylation site"
FT 488..492
FT /note= "N-glycosylation site"
FT 531..537
FT /note= "N-myristoylation site"
FT 579..583
FT /note= "Casein Kinase II phosphorylation site"
FT 606..610
FT /note= "N-glycosylation site"
FT 620..624
FT /note= "Casein Kinase II phosphorylation site"
XX
PN WO200015797-A2.
XX
XX 23-MAR-2000.
XX PD
XX 15-SEP-1999; 99WO-US21547.
XX PR
XX 17-SEP-1998; 98US-0100858.
XX PR 17-SEP-1998; 98WO-US19437.
XX (GETH) GENENTECH INC.
XX PA Fong S, Goddard A, Gurney AL, Tumas D, Wood WI;
XX PI WPI; 2000-271435/23.
XX DR N-PSDB; AA252207.
XX PT Composition for treatment and diagnosis of immune related diseases e.g.
PT Grave's disease comprises a PRO245, PRO217, PRO301, PRO266, PRO335,
PT PRO331 or PRO326 polypeptide or its agonists or antagonists (preferably
PT antibodies) -
XX
XX Example 1; Fig 14; 201pp; English.
XX
XX The present sequence is the human protein PRO331, encoded by UNQ292 cDNA,
XX designated as clone DNA40981. It is isolated from human foetal brain
XX tissue. Portions of PRO331 has homology to the IIG-1 protein.
XX It enhances or suppresses the infiltration of inflammatory cells into
XX tissues, proliferation of T-lymphocytes and modulates the immune
XX response. This sequence is useful for treatment of immune related
XX disorders, like SLE, rheumatoid/juvenile arthritis, spondyloarthropathy,
XX systemic sclerosis (scleroderma), idiopathic inflammatory myopathies such
XX as dermatomyositis, Sjogren's syndrome, systemic vasculitis, sarcoidosis,
XX autoimmune haemolytic anaemia, thrombocytopenia, thyroiditis e.g. Grave's
XX disease, diabetes mellitus, immune-mediated renal disease e.g.
XX glomerulonephritis, demyelinating diseases such as multiple sclerosis and
XX Guillain-Barre syndrome, hepatobiliary diseases like hepatitis and
XX primary biliary cirrhosis, inflammatory and fibrotic lung diseases such
XX as inflammatory bowel disease (e.g. Crohn's disease), autoimmune or
XX immune-mediated skin diseases such as psoriasis, allergies like asthma,
XX immunological diseases of the lungs such as eosinophilic pneumonia and
XX transplantation associated diseases such as graft-versus-host-disease.
XX
XX Sequence 640 AA:
XX
Query Match 100.0%; Score 2524; DB 21; Length 640;
Best Local Similarity 100.0%; Pred No. 2.4e-177;
Matches 478; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QTCPSVCSNQSFKVICVKNLRNRPDGISTRNLNLHENOIQIIVKNSFKHLRLEI 60
Db 45 qtcpsvcscnqsfkvcivcrknirevpgdgistrrllnlhcnqikvnsfkhlrlei 104
QY 61 LQLSRNHRTIEIGAFNGLANLNTLELFDNRLTIPNGAFVYLSKLELNRNPISPI 120

Db 105 lqlsrnhrtieigafnglanlntlelfdnrlttipngafvyskklwlrnnpiesip 164
QY 121 SYAFNRTPSLRRDLGELKRLSYISEGAFGLSNLRYLNAMCNLREIPNLTPLIKLDEL 180
Db 165 syafnrtpslrrdlgelkrlsyisegafeglsnrylnamcnlreipnltplikidel 224
QY 181 DLGNHLSAIRPGSFQGLMHLQKLMWTSQIQVIERNAFDNLQSLVEINLAHNLTLLPH 240
Db 225 dlsgnhlsairpgsfqglmhlqklwmqlsqvqviernafdnlsqveinlahnltllph 284
QY 241 DLFTPLHLRIHLHNPWNCNDILWLSWIKMDAPSNATACCARCNTPPNLKGRYIGEL 300
Db 285 dlftplhlrlhlhnpwncndilwlswwikmdapsntaccarcntppnlkgriygel 344
QY 301 DQNYFTCYAPVIVEPPADLNVTGMAELKCRASSTLSVSWITPNCVTMTHGAYKVRIA 360
Db 345 dqnyftcyapviveppadlnvtgmaelkcraststsvswitpncvtmthgaykvria 404
QY 361 VLSGTLNFTNVTVDGMYTCMVNSVGNNTASATLNVTAAATTPFSYFSTVTMETEP 420
Db 405 vlsdgtlnftnvtvdgmytcmvnsvgnttasatlnvtaatttfsyfstvtmetep 464
QY 421 SQDEARTDNNVGPVVDHETNVTLSLPQSTRTEKFTTIPVTDINSIGPIDV 478
Db 465 sqdeartdnnvgpvdvhdetnvtlslpqstrstekfttippvtdinsigpidv 522
RESULT 6
AAU12355
ID AAU12355 standard; Protein: 640 AA.
XX
AC AAU12355;
XX
DT 24-OCT-2001 (first entry)
XX
DE Human PRO331 polypeptide sequence.
XX
KW Human secretory and transmembrane; PRO: mammalian; cancer; lung;
KW breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;
KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
KW adipocyte; A-peptide; factor VIIA; gene therapy.
XX
OS Homo sapiens.
XX
PN WO20010466-A2.
XX
PD 07-JUN-2001.
XX
PF 01-DEC-2000; 2000WO-US32678.
XX
PR 01-DEC-1999; 99WO-US28301.
PR 01-DEC-1999; 99WO-US28634.
PR 02-DEC-1999; 99WO-US28551.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.
PR 09-DEC-1999; 99US-0170262.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 30-DEC-1999; 99WO-US31243.
PR 06-JAN-2000; 2000WO-US00277.
PR 06-JAN-2000; 2000WO-US00376.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US04914.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 20-MAR-2000; 2000WO-US07377.
PR 21-MAR-2000; 2000WO-US07532.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.

PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 10-NOV-2000; 2000WO-US30873.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX
DR WPI; 2001-408281/43.
DR N-PSDB; AAS21427.
XX
PT Isolated, secretory and transmembrane PRO polypeptide used to detect
other PRO polypeptides, link bioactive molecules to cells expressing
PRO polypeptides, and detect the presence of mammalian tumours e.g.
lung, breast, prostate, cervical
XX
PS Claim 12; Fig 368; 813pp; English.
XX
CC AAU12172-AAU12446 represent novel human secretory and transmembrane
PRO polypeptides. The PRO polypeptides are useful to detect other
PRO polypeptides, to link bioactive molecules to cells expressing
PRO polypeptides, to modulate biological activities of cells expressing
PRO polypeptides, and to detect the presence of mammalian lung, colon,
breast, prostate, rectal, cervical or liver tumours by comparing PRO
polypeptide expression in a cell sample to that in a control sample.
CC Some of the 275 sequences are also useful to stimulate the release of
tumour necrosis factor-alpha (TNF-alpha) from human blood, the
CC proliferation or differentiation of chondrocytes, the proliferation or
gene expression in pericyte cells, the release of proteoglycans from
cartilage, the proliferation of inner ear utricular supporting cells or
of T-lymphocytes, the release of a cytokine from peripheral blood
CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of
the PRO polypeptides may modulate glucose or free fatty acid uptake by
CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
to factor VIIA. The PRO polypeptides can be used in assays to identify
CC molecules involved in binding interactions. The polynucleotides encoding
PRO polypeptides can be used to generate probes, antisense RNA/DNA,
CC transgenic or knock out animals and can be used in gene therapy.
XX
SQ Sequence 640 AA;
Query Match 100.0%; Score 2524; DB 22; Length 640;
Best Local Similarity 100.0%; Pred. No. 2.4e-177; Indels 0; Gaps 0;
Matches 478; Conservative 0; Mismatches 0;
QY 1 QTCPSVCSQNSQFVKVCRKRLREVPDGIETNTRRLNLHENQIQIKVNSFKHLRLEI 60
Db 45 qtcpsvcscsqnsqfsvkvcvcrknlrevpdgistntrlnlhengqiqikvnsfkhlrllei 104
QY 61 LQLSRNHRTIEIGAFNGLANLTLELFDNRLTTPNGAFVYLSKELWLRNNPISIP 120
Db 105 lqlsrnhrtieigafnglanlntlelfdnrlttipngafvyskklwlrnnpiesip 164
QY 121 SYAFNRTPSLRRDLGELKRLSYISEGAFGLSNLRYLNAMCNLREIPNLTPLIKLDEL 180
Db 165 syafnrtpslrrdlgelkrlsyisegafeglsnrylnamcnlreipnltplikidel 224
QY 181 DLGNHLSAIRPGSFQGLMHLQKLMWTSQIQVIERNAFDNLQSLVEINLAHNLTLLPH 240
Db 225 dlsgnhlsairpgsfqglmhlqklwmqlsqvqviernafdnlsqveinlahnltllph 284
QY 241 DLFTPLHLRIHLHNPWNCNDILWLSWIKMDAPSNATACCARCNTPPNLKGRYIGEL 300
Db 285 dlftplhlrlhlhnpwncndilwlswwikmdapsntaccarcntppnlkgriygel 344
QY 301 DQNYFTCYAPVIVEPPADLNVTGMAELKCRASSTLSVSWITPNCVTMTHGAYKVRIA 360
Db 345 dqnyftcyapviveppadlnvtgmaelkcraststsvswitpncvtmthgaykvria 404

QY 361 VLSGTLNFTNVTQDGMVTCMVSNSVGNVTASATLNVAATTPRSYFSTVTETMEP 420
 Db 405 vlsdgtlnfnvqvdtgmytcmvsnvsgntasatlnvtaatttptfsyfstvtetmep 464
 QY 421 SQDEARTDNNVGPVVDWETNVTSLPQSTRSTEKFTTPVTDSNGIPGIDEV 478
 Db 465 sqdeartdnnvqptpvvdwettntvtsltpqstrstekfttpvtdnsngipgidev 522

RESULT 7
 AAU00826
 ID AAU00826 standard; Protein; 640 AA.
 XX
 AC AAU00826:
 XX
 DT 04-JUL-2001 (first entry)
 XX
 DE Human immune response protein PRO331 (UNQ292).
 XX
 KW Human; PRO331; UNQ292; immune response; osteoarthritis;
 KW systemic lupus erythematosus; rheumatoid arthritis; systemic sclerosis;
 KW juvenile chronic arthritis; spondyloarthropathy; Sjogren's syndrome;
 KW idiopathic inflammatory myopathy; polymyositis; systemic vasculitis;
 KW sarcoidosis; autoimmune haemolytic anaemia; immune pancytopenia;
 KW autoimmune thrombocytopaenia; idiopathic thrombocytopaenic purpura;
 KW thyroiditis; Grave's disease; Hashimoto's thyroiditis;
 KW diabetes mellitus; glomerulonephritis; demyelinating disease;
 KW multiple sclerosis; Guillain-Barre syndrome; hepatobiliary disease;
 KW chronic inflammatory demyelinating polyneuropathy; infectious hepatitis;
 KW auto immune chronic active hepatitis; primary biliary cirrhosis;
 KW granulomatous hepatitis; sclerosing cholangitis; ulcerative colitis;
 KW inflammatory bowel disease; Crohn's disease; Whipple's disease;
 KW erythema multiforme; psoriasis; asthma; allergic rhinitis; urticaria;
 KW food hypersensitivity; eosinophilic pneumonia; graft rejection;
 KW idiopathic pulmonary fibrosis; graft-versus-host-disease; immunogen;
 KW antibody.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT Peptide
 FT 1..44
 FT /label= Signal_peptide
 FT Modified-site 40..46 /note= "Glycine is N-myristoylated"
 FT Protein 45..640
 FT /label= Mature_PRO331
 FT Modified-site 73..79
 FT /note= "Glycine is N-myristoylated"
 FT Modified-site 118..124
 FT /note= "Glycine at 118 is N-myristoylated"
 FT Region 183..187
 FT /label= Phosphorylation_site
 FT /note= "CAMP/CGMP dependent protein kinase phosphorylation site"
 FT Modified-site 191..197
 FT /note= "Glycine at 191 is N-myristoylated"
 FT Modified-site 228..234
 FT /note= "Glycine is N-myristoylated"
 FT Modified-site 237..243
 FT /note= "Glycine is N-myristoylated"
 FT Modified-site 278..282
 FT /note= "Asn is N-glycosylated"
 FT Modified-site 364..368
 FT /note= "Asn is N-glycosylated"
 FT Modified-site 390..394
 FT /note= "Asn is N-glycosylated"
 FT Modified-site 391..397
 FT /note= "Glycine at 391 is N-myristoylated"
 FT Modified-site 412..416
 FT /note= "Asn is N-glycosylated"
 FT Modified-site 415..419
 FT /note= "Asn is N-glycosylated"
 FT Modified-site 422..428

FT Modified-site 433..439 /note= "Glycine is N-myristoylated"
 FT /note= "Glycine is N-myristoylated"
 FT Modified-site 434..438
 FT /note= "Asn is N-glycosylated"
 FT Modified-site 442..446 /note= "Asn is N-glycosylated"
 FT Modified-site 488..492 /note= "Asn is N-glycosylated"
 FT Domain 528..543 /label= Transmembrane_domain
 FT Modified-site 531..537 /note= "Glycine is N-myristoylated"
 FT Modified-site 606..610 /note= "Asn is N-glycosylated"
 XX
 PN WO200119991-A1.
 XX
 PD 22-MAR-2001.
 XX
 PF 20-MAR-2000; 2000WO-US07377.
 XX
 PR 15-SEP-1999; 99WO-US21547.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Fong S, Goddard A, Gurney AL, Hillan KJ, Tumas D, Wood WI;
 XX WPI; 2001-226823/23.
 DR N-PSDB; AAS00162.
 XX
 PT Composition for diagnosing and treating immune related diseases, e.g.
 PT rheumatoid arthritis and diabetes mellitus, comprises a PRO
 PT polypeptide, agonist, antagonist or fragment -
 XX
 PS Claim 31; Fig 12; 138pp; English.
 XX
 CC The sequence represents Human PRO331 (UNQ292), a protein involved in
 CC the immune response. PRO polypeptides, and (ant)agonists to them, are
 CC used in compositions for modulating infiltration of inflammatory cells
 CC into a tissue, modulating an immune response and modulating proliferation
 CC of T-lymphocytes in response to an antigen. Immune related diseases can
 CC be treated with the compositions, such as, systemic lupus erythematosus,
 CC rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis,
 CC spondyloarthropathies, systemic sclerosis, idiopathic inflammatory
 CC myopathies (e.g. polymyositis), Sjogren's syndrome, systemic vasculitis,
 CC sarcoidosis, autoimmune haemolytic anaemia (e.g. immune pancytopenia),
 CC autoimmune thrombocytopaenia (e.g. idiopathic thrombocytopaenic purpura),
 CC thyroiditis (e.g. Grave's disease, Hashimoto's thyroiditis), diabetes
 CC mellitus, immune-mediated renal disease (e.g. glomerulonephritis),
 CC demyelinating diseases of the central and peripheral nervous systems e.g.
 CC multiple sclerosis or Guillain-Barre syndrome, and chronic inflammatory
 CC demyelinating polyneuropathy, hepatobiliary diseases such as infectious
 CC hepatitis (hepatitis A, B, C, D, E and other non-hepatotropic viruses),
 CC auto immune chronic active hepatitis, primary biliary cirrhosis,
 CC granulomatous hepatitis, and sclerosing cholangitis, inflammatory bowel
 CC disease (ulcerative colitis, Crohn's disease and Whipple's disease),
 CC autoimmune or immune-mediated skin diseases (e.g. erythema
 CC multiforme and psoriasis), asthma, allergic rhinitis, urticaria,
 CC food hypersensitivity, immunologic diseases of the lung such as
 CC eosinophilic pneumonias, idiopathic pulmonary fibrosis, transplantation
 CC associated diseases including graft-versus-host-disease and graft
 CC rejection. PRO polypeptides can be used to diagnose immune related
 CC diseases, to identify inhibitors, and to stimulate the proliferation of
 CC T lymphocytes. Anti-PRO antibodies can be used to detect PRO and
 CC in diagnosis. PRO polypeptides, antibodies and (ant)agonists can be used
 CC in rational drug design.
 XX
 SQ Sequence 640 AA;

Query Match 100.0%; Score 2524; DB 22; Length 640;
 Best Local Similarity 100.0%; Pred. No. 2.4e-177;

Matches 478; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QTCPSVCSNQFSKVICVRKNLEVPDGI STNRLNLHENOIQIIVKNSFKHLRLEI 60
 DB 45 qtcpsvcsnqfskvicvrknlevpdgi stnrlnlhenoqiiikvnsfkhlrllei 104
 QY 61 LQSRNHRTIEIGAFNGLANLTLEFDNRLLTTPNGAFVYLSKRLKELWLRNPIESIP 120
 DB 105 lqsrnhrtieigafnglanltlefdnrlttipngafvylsklkelwlrnnpiesip 164
 QY 121 SYAFNRIPSLRRDLGELKRLSYISEGAFGLSNLRYLNLMCNLREIPNLTPLIKDEL 180
 DB 165 syafnrpslrrldgelkrlsyisegafeglsnrylnlmcnrlreipnltplikidel 224
 QY 181 DLSGNHLSAIRPGSFQGLMHQKLMWIOSQIQTIERNAFNLQSLVEINLAHNLTLLPH 240
 DB 225 dlsghnlsairpgsfqglmhqklmwioqqlvieraafnldqslveinlahnltllph 284
 QY 241 DLFTPLHHLERIHLLHNPWNCNDILWLSMWIKDMAFNTACCRCNTPPNKGRIYIGEL 300
 DB 285 dlftplhhlrlihllhnpwncndilwlswwikdmapsntaccarcntppnkgryigel 344
 QY 301 DQNYFTCYAPVIVEPPADLVNTEGMAAEKCRASSTLSVSWITPNGTVMTHGAYKVRIA 360
 DB 345 dqnyftcyapviveppadlvntegmaaelkcrastsitsvswitpngtmthgaykvria 404
 QY 361 VLSGDTLNFTNVTQDTGMVTCWVSNVGNVTASATLNVTAAATTPFSYFSTVVTMEP 420
 DB 405 vlsdgtlnftnvtqdtgmytcwvsnvsgntasatlntvtaatttppsystvvtmep 464
 QY 421 SQDEARTDNNVGPVVDWETTNVTTSLTPQSTRSTEKFTTIPVTDINSIGPIDEV 478
 DB 465 sqdeartdnnvgpvdvwdettnvttsltppqstrstekfttipvtdinsigpidev 522

RESULT 8
 AAB80262
 ID AAB80262 standard; Protein: 640 AA.
 XX
 AC AAB80262;
 XX
 DT 24-APR-2001 (first entry)
 XX
 DE Human PRO331 protein.
 XX
 KW Human; PRO; dermatological; antipsoriatic; cytostatic; antiinflammatory;
 KW antiparkinsonian nootropic; neuroprotective; vulnery; cardiac;
 KW antiangiogenic; vasotropic; antiasthmatic; antirheumatic; cancer;
 KW antiarthritic; antilinfertility; antidiabetic; antiviral; diabetes;
 KW ophthalmological; gene therapy; skin disease; gastrointestinal disorder;
 KW ischaemia; inflammation.
 XX
 OS Homo sapiens.
 XX
 PN WO200104311-A1.
 XX
 PD 18-JAN-2001.
 XX
 PF 22-FEB-2000; 2000WO-US04414.
 XX
 PR 07-JUL-1999; 99US-0143048.
 PR 26-JUL-1999; 99US-0145698.
 PR 28-JUL-1999; 99US-0146222.
 PR 08-SEP-1999; 99WO-US20594.
 PR 13-SEP-1999; 99WO-US20944.
 PR 15-SEP-1999; 99WO-US21090.
 PR 15-SEP-1999; 99WO-US21547.
 PR 05-OCT-1999; 99WO-US23089.
 PR 29-NOV-1999; 99WO-US28214.
 PR 30-NOV-1999; 99WO-US28313.
 PR 16-DEC-1999; 99WO-US30095.
 PR 20-DEC-1999; 99WO-US30911.
 PR 20-DEC-1999; 99WO-US30999.

PR 05-JAN-2000; 99WO-US00219.
 XX (GETH) GENENTECH INC.
 PA Ashkenazi AJ, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
 XX Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
 PI Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ, Kljavin IG;
 PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
 PI Williams PM, Wood WI;
 XX WPI: 2001-081051/09.
 DR N-PSDB; AAF72423.
 XX
 XX Sixty one nucleic acids encoding PRO polypeptides which are useful in
 PT the treatment of skin diseases (e.g. psoriasis), cancers (e.g. lung
 PT squamous cell carcinoma) and neurodegenerative diseases (e.g.
 PT Alzheimer's disease) -
 XX
 XX Claim 1: Fig 104; 393pp; English.
 PS
 PS The present sequence is one of sixty one novel secreted and
 CC transmembrane PRO polypeptides. The PRO polypeptides are
 CC useful for treating skin diseases (e.g. psoriasis), cancers (e.g. lung
 CC squamous cell carcinoma), gastrointestinal disorders (e.g.
 CC enterocolitis), neurodegenerative diseases (e.g. Alzheimer's disease,
 CC Parkinson's disease), wound repair, cardiovascular disorders (e.g.
 CC endometrial bleeding angiogenesis, ischaemias such as coronary
 CC ischaemia, atherosclerosis), inflammatory disorders (e.g. asthma,
 CC rheumatoid arthritis, multiple sclerosis), infertility, AIDS and
 CC diabetes and retinal disorders such as retinitis pigmentosum.
 CC The PRO nucleic acids have applications in molecular biology, including
 CC use as hybridization probes, and in chromosome and gene mapping.
 XX
 XX Sequence 640 AA;

Query Match 100.0%; Score 2524; DB 22; Length 640;
 Best Local Similarity 100.0%; Pred. No. 2.4e-177;
 Matches 478; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QTCPSVCSNQFSKVICVRKNLEVPDGI STNRLNLHENOIQIIVKNSFKHLRLEI 60
 DB 45 qtcpsvcsnqfskvicvrknlevpdgi stnrlnlhenoqiiikvnsfkhlrllei 104
 QY 61 LQSRNHRTIEIGAFNGLANLTLEFDNRLLTTPNGAFVYLSKRLKELWLRNPIESIP 120
 DB 105 lqsrnhrtieigafnglanltlefdnrlttipngafvylsklkelwlrnnpiesip 164
 QY 121 SYAFNRIPSLRRDLGELKRLSYISEGAFGLSNLRYLNLMCNLREIPNLTPLIKDEL 180
 DB 165 syafnrpslrrldgelkrlsyisegafeglsnrylnlmcnrlreipnltplikidel 224
 QY 181 DLSGNHLSAIRPGSFQGLMHQKLMWIOSQIQTIERNAFNLQSLVEINLAHNLTLLPH 240
 DB 225 dlsghnlsairpgsfqglmhqklmwioqqlvieraafnldqslveinlahnltllph 284
 QY 241 DLFTPLHHLERIHLLHNPWNCNDILWLSMWIKDMAFNTACCRCNTPPNKGRIYIGEL 300
 DB 285 dlftplhhlrlihllhnpwncndilwlswwikdmapsntaccarcntppnkgryigel 344
 QY 301 DQNYFTCYAPVIVEPPADLVNTEGMAAEKCRASSTLSVSWITPNGTVMTHGAYKVRIA 360
 DB 345 dqnyftcyapviveppadlvntegmaaelkcrastsitsvswitpngtmthgaykvria 404
 QY 361 VLSGDTLNFTNVTQDTGMVTCWVSNVGNVTASATLNVTAAATTPFSYFSTVVTMEP 420
 DB 405 vlsdgtlnftnvtqdtgmytcwvsnvsgntasatlntvtaatttppsystvvtmep 464
 QY 421 SQDEARTDNNVGPVVDWETTNVTTSLTPQSTRSTEKFTTIPVTDINSIGPIDEV 478
 DB 465 sqdeartdnnvgpvdvwdettnvttsltppqstrstekfttipvtdinsigpidev 522

RESULT 9
AAB65292
ID AAB65292 standard; Protein: 640 AA.
XX AC AAB65292;
XX DT 02-APR-2001 (first entry)
XX DE Human PRO331 protein sequence SEQ ID NO:501.
XX DE
XX KW Human; secreted and transmembrane protein; PRO; cytostatic;
KW cell death; cancer; chromosomal mapping; gene mapping; tissue typing;
KW diagnostic assay.
XX OS Homo sapiens.
XX PN W0200073454-A1.
XX PD
XX PF 07-DEC-2000.
XX PF 30-MAR-2000; 2000WO-US08439.
XX PR 02-JUN-1999; 99WO-US12252.
XX PR 23-JUN-1999; 99US-0141037.
XX PR 07-JUL-1999; 99US-0143048.
XX PR 20-JUL-1999; 99US-0144758.
XX PR 26-JUL-1999; 99US-0145698.
XX PR 28-JUL-1999; 99US-0146222.
XX PR 17-AUG-1999; 99US-0149396.
XX PR 15-SEP-1999; 99WO-US21090.
XX PR 15-SEP-1999; 99WO-US21547.
XX PR 08-OCT-1999; 99US-0158663.
XX PR 30-NOV-1999; 99WO-US28313.
XX PR 01-DEC-1999; 99WO-US28301.
XX PR 16-DEC-1999; 99WO-US30095.
XX PR 20-DEC-1999; 99WO-US30911.
XX PR 03-JAN-2000; 2000WO-US00219.
XX PR 06-JAN-2000; 2000WO-US00376.
XX PR 11-FEB-2000; 2000WO-US03565.
XX PR 18-FEB-2000; 2000WO-US04341.
XX PR 22-FEB-2000; 2000WO-US04414.
XX PR 24-FEB-2000; 2000WO-US04914.
XX PR 24-FEB-2000; 2000WO-US05004.
XX PR 02-MAR-2000; 2000WO-US05841.
XX PR 15-MAR-2000; 2000WO-US06884.
XX PR 20-MAR-2000; 2000WO-US07377.
XX PR
XX (GETH) GENENTECH INC.
XX PI Ashkenazi AJ, Baker KP, Botstein D, Desnovers L, Eaton DL;
PI Ferrara N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi CJ, Gurney AL, Kijavini IJ, Napier MA, Pan J, Paoni NF;
PI Roy MA, Stewart RA, Tamas D, Watanabe CK, Williams PM, Wood WI;
PI Zhang Z;
XX WPi: 2001-032160/04.
XX N-PSDB; AAF44261.
XX
XX PRO polynucleotides used to produce polypeptides used to target
PT bioactive molecules such as toxins, radiolabels or antibodies, to
PT specific cells, to cause targeted cell death -
XX
XX Claim 12; Fig 314; 935pp; English.
XX
XX The present invention describes human secreted and transmembrane PRO
CC proteins. The PRO proteins have cytostatic activity. The PRO proteins
CC can be used for targeted delivery of bioactive molecules, such as
CC toxins, radiolabels or antibodies, that cause cell death. PRO nucleotide
CC sequences, and their fragments, can be used as hybridisation probes, in
CC chromosomal and gene mapping, and in the generation of anti-sense RNA
CC and DNA. They may also be used to produce transgenic animals which are
CC used to develop and screen therapeutically useful reagents. The PRO
CC nucleotide and protein sequence can be used for tissue typing and in

CC treating cancer. Anti-PRO antibodies can be used in diagnostic assays.
CC AAF44270 to AAF44470 represent PCR primers and hybridisation probes used
CC in the isolation of human PRO sequences. AAF44087 to AAF44269 and
CC AAB65154 to AAB65300 represent human PRO polynucleotide and protein
CC sequences given in the exemplification of the present invention.
XX
SQ Sequence 640 AA;
Query Match 100.0%; Score 2524; DB 22; Length 640;
Best Local Similarity 100.0%; Pred. No. 2.4e-177;
Matches 478; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QTCPSVCSNQSFSKVCVRKNLREVDPGISTNRLNLHENQIQIKVNSFKHLRLEI 60
DB 45 qtcpsvcsnqsfkvcvrknirevpdgiistnrlnlhenqikvnsfkhlrlei 104
QY 61 LQLSRNHIRTIEIGAFNGLANLTLFELFDNRLLTIPNGAFVYLSKLKELWLRNPPIESIP 120
DB 105 lqlsrnhirtieigafnglanlntlelfdnrlltipngafvylskklwlrnpiesip 164
QY 121 SYAFNRIPSLRLDLGELKRLSYISEGAFGLSNLRYNLNLCMLNREIPNLTPLIKLDEL 180
DB 165 syafnrpslrrldlgelkrlsyisegafeglsnlrynlncmlnreipnltplikidel 224
QY 181 DLGNHLSAIRPGSFOGLMHLQKLMWIOSQIQVIERAFNQLQSLVEINLAHNNLTLLPH 240
DB 225 dlsgnhlsairpgsfqglmhlqkwmqsgqlqvierafnqlqslveinlahnnltllph 284
QY 241 DLFTPLHHLERILHHPNWCNCDILWLSWIKDMAPSNACCACRNTPPNPKGRYIGEL 300
DB 285 dlftplhhlerrilhnpwncndilwlswwikdmapsntaccacrcntppnalkgryigel 344
QY 301 DQNYFTCYAPVIVEPPADLNVTGMAAELKCRASSTLSYSWITPNCGTVMTHGAYKVRIA 360
DB 345 dqnyftcyapviveppadlnvtgmaaelkcrasstlsyswitsvtpngtvmthgaykvria 404
QY 361 VLSGTLNFTNVTVDTCMTVMYSNSVGNFTASATLNVTAAATTPFSYFSTVVTMEP 420
DB 405 vlsogtlnftnvtvdtcmtvmysnsvgnftasatlnvtaatttfsyfstvtvetmep 464
QY 421 SQDEARTDNNVGTPPVVDWETTNVTTSLTPQSTRSTEKFTPIPVTDINSIGIDEV 478
DB 465 sqdeartdnnvgtpvvdwettntvttsltpqstrstekfttipvtdinsigpidev 522
RESULT 10
AAB53089
ID AAB53089 standard; Protein: 640 AA.
XX AC AAB53089;
XX DT 28-FEB-2001 (first entry)
XX
XX Human angiogenesis-associated protein PRO331, SEQ ID NO:137.
XX DE
XX KW Human; angiogenesis-associated protein; PRO; endothelial cell growth;
KW cardiac hypertrophy; cardiovascular disorder; endothelial disorder;
KW angiogenic disorder; atherosclerosis; osteoporosis; hypertension;
KW myocardial infarction; diabetic retinopathy; rheumatoid arthritis;
KW Crohn's disease; psoriasis; endometriosis; ulcer; wound healing; cancer;
KW Alzheimer's disease; Huntington's disease; stroke; drug screening;
KW gene therapy; transgenic animal.
XX
XX Homo sapiens.
XX OS
XX PN W0200053753-A2.
XX
XX PD 14-SEP-2000.
XX
XX PF 05-JAN-2000; 2000WO-US00219.
XX PF 08-MAR-1999; 99WO-US05028.
XX PR

PR 12-MAR-1999; 99US-0123957.
 PR 14-MAY-1999; 99US-0134287.
 PR 02-JUN-1999; 99WO-US12252.
 PR 23-JUN-1999; 99US-0141037.
 PR 20-JUL-1999; 99US-0144758.
 PR 26-JUL-1999; 99US-0145698.
 PR 01-SEP-1999; 99WO-US20111.
 PR 08-SEP-1999; 99WO-US20594.
 PR 13-SEP-1999; 99WO-US21090.
 PR 15-SEP-1999; 99WO-US21347.
 PR 03-OCT-1999; 99WO-US23089.
 PR 30-NOV-1999; 99WO-US28313.
 PR 30-NOV-1999; 99WO-US28409.
 PR 02-DEC-1999; 99WO-US28564.
 PR 02-DEC-1999; 99WO-US28565.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Ashkenazi AJ, Baker KP, Ferrara N, Gerber H, Goddard A;
 PI Godowski PJ, Gurney AL, Hillan KJ, Kuo SS, Mark MR, Marsters SA;
 PI Paoni NF, Pitti RM, Watanabe CK, Williams PM, Wood WI;
 XX
 DR WPI; 2001-090793/10.
 DR N-PSDB; AAC97475.
 XX
 PT New isolated nucleic acid for producing a PRO polypeptide, analyzing
 PT genetic disorders and treating cardiovascular, endothelial or
 PT angiogenic disorders, such as atherosclerosis, wounds or cancer -
 XX
 PS Claim 69; Fig 52; 293pp; English.
 XX
 CC The invention relates to novel human angiogenesis-associated proteins
 CC designated PRO proteins (AAB53064-B53097), and to nucleic acids encoding
 CC PRO proteins. The invention also relates to vectors and host cells
 CC comprising a PRO nucleic acid, the recombinant production of a PRO
 CC protein, PRO antibodies specific for a PRO protein, fusion proteins
 CC comprising a PRO protein, agonists or antagonists of a PRO protein, and
 CC compounds which inhibit the expression of a PRO gene. The invention
 CC additionally encompasses methods of identifying modulators of PRO
 CC expression or activity; diagnosing a cardiovascular, endothelial or
 CC angiogenic disorder, or a susceptibility to such a disorder by detecting
 CC mutations in a PRO gene, or the expression level of a PRO gene within a
 CC particular tissue; treating a cardiovascular, endothelial or angiogenic
 CC disorder via the administration of a PRO protein, PRO nucleic acid, or
 CC PRO agonist or antagonist; a retroviral gene therapy vector comprising a
 CC PRO nucleic acid; and methods of inhibiting or stimulating endothelial
 CC cell growth, cardiac hypertrophy or PRO-induced angiogenesis via the
 CC administration of a PRO protein, or an agonist or antagonist thereof.
 CC PRO nucleic acids, PRO proteins, or an agonist or antagonist thereof.
 CC agonists and PRO antagonists may be used as therapeutic agents to treat
 CC cardiovascular, endothelial or angiogenic disorders, such as
 CC atherosclerosis, osteoporosis, myocardial infarction, hypertension,
 CC diabetic retinopathy, rheumatoid arthritis, Crohn's disease, psoriasis,
 CC endometriosis, ulcers, wounds, cancer, Alzheimer's disease, Huntington's
 CC disease, or stroke. PRO nucleic acids are additionally useful in the
 CC recombinant production of PRO proteins, as hybridisation probes to
 CC screen libraries to isolate cDNAs with sequence identity to PRO proteins,
 CC to map genes encoding PRO proteins, to analyse genetic disorders, and in
 CC gene therapy. PRO nucleic acids can also be used to produce transgenic
 CC animals useful for the development and screening of potential
 CC therapeutic agents. The present sequence represents a PRO protein of the
 CC invention.
 XX
 SQ Sequence 640 AA;

Query Match 100.0%; Score 2524; DB 22; Length 640;
 Best Local Similarity 100.0%; Pred. No. 2.4e-177;
 Matches 478; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QTCPSVCSCSNQFQSKVTCVRKNLREVPDGTSTNRLNLHNEHQIQLKVNSEKHLRLEI 60
 DB 45 qtcpsvcscsnqfqskskvcvrknrlrevpdgstnrlnlhnehqqlkvnsefknrlhle 104

QY 61 LQLSRNHRTIEIGAFNGLANLNTLEFDNRLTTPNGAFVYLSKLKELWLNPNIESIP 120
 DB 105 lqlsrnhrtieigafnglanlntlefdnrlttpngafvysklkelwlnpniesip 164
 QY 121 SYAFNRIPSLRRDLGELKRLSYISEGAFGLSNRLRYINLAMCNLREIPNLTPILKIDEL 180
 DB 165 syafnrripslrrdlgelkrlsyisegafeglsnrlryniamcnlreipnltpikidel 224
 QY 181 DLSGNHLSATRPSFGQFMHLQKLMWIOSQIQVIERNAFDNLQSLVEINLAHNNLTLLPH 240
 DB 225 dlsghnhsairpsfgqfmhqlkqlwmigsgqiviernafdnlgslveinlahnnltllph 284
 QY 241 DLFTPLHLERILHHPNWCNCDILWSWIKDMAPSNATACCARCNTPPNKGKRYIGEL 300
 DB 285 dlftplhlhlerilhnpwncncdilwswikdmapsntaccarcnppnkgkryigel 344
 QY 301 DQNYFTCYAPVIVEPPADLNVTGMAAELKCRASSTLSVSWITPNGVTMTHGAYKVRIA 360
 DB 345 dqnyftcyapviveppadlnvtgmaaelkcrasstlsvswitpngvtmthgaykvrria 404
 QY 361 VLSGDTLNFNTVQDTGMYTCMVSNVGNVTASATLNVTAAATTPPSYFSTVIVTMEP 420
 DB 405 vlsdgtlnfntvqdtgmytcmvsvngvntasatlnvtaatttppsfystvtvmetep 464
 QY 421 SQDEARTDNNVGPTPVVDWETTNTVTTSLTPQSTRSTEKTTIPVTDINSIGIPGIDEV 478
 DB 465 sqdeartdnnvgptpvvdwettntvttsltpqstrstekttipvtdinsigipgidev 522

RESULT 11
 AAEL3006
 ID AAEL3006 standard; Protein; 713 AA.
 XX
 AC AAEL3006;
 XX
 DT 28-JAN-2002 (first entry)
 XX
 DE Human leucine-rich repeat (LRR) family member protein.
 XX
 KW Human; leucine-rich repeat; LRR; 31939 protein; therapy;
 KW cell proliferation; differentiation disorder; cancer; neuronal disorder;
 KW neurological disorder; demyelinating disease; multiple sclerosis;
 KW degenerative disease; Alzheimer's disease; Huntington's disease;
 KW spinocerebellar degeneration; nervous system; bone disorder;
 KW osteoporosis; immune disorder; rheumatoid arthritis; diabetes mellitus;
 KW cardiovascular disorder; liver disorder; viral disease; pain;
 KW metabolic disorder; chromosomal mapping; tissue typing; forensic biology;
 KW cytostatic; nootropic; neuroprotective; anticonvulsant; osteopathic;
 KW antirheumatic; antiarthritic; virucide; analgesic.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..38
 FT /label= Signal_peptide
 FT Protein 39..713
 FT /note= "Mature human 31939 protein"
 FT Domain 56..85
 FT /label= N-terminal_LRR_domain
 FT Domain 87..110
 FT /label= LRR_domain
 FT Domain 111..134
 FT /label= LRR_domain
 FT Domain 135..158
 FT /label= LRR_domain
 FT Domain 159..182
 FT /label= LRR_domain
 FT Domain 183..207
 FT /label= LRR_domain
 FT Domain 208..229
 FT /label= LRR_domain
 FT Domain 230..253

PT adult brain, human fetal brain, human fetal kidney, and human adult
PT blood cDNA libraries
XX
PS Claim 1; Page 97-99; 122pp; English.
XX
CC The present sequence is the cc359_4 secreted protein encoded by the cDNA
CC clone cc359_4. cc359_4 was isolated from a human adult brain cDNA library
CC using methods specific for secreted protein cDNAs. The leader sequence or
CC signal peptide acts as a transmembrane domain due to its hydrophobic
CC nature. The TopPredII computer program predicts five potential
CC transmembrane domains centered around amino acids 20, 410, 490, 530 and
CC 590. This protein has a leucine zipper motif. The polynucleotide and
CC protein may effect nutritional activity, cytokine and cell proliferation,
CC immune stimulation or suppression, hematopoiesis regulation, tissue
CC growth, tumour inhibition etc.
XX
SQ Sequence 653 AA;

Query Match 62.7%; Score 1581.5; DB 20; Length 653;
Best Local Similarity 60.9%; Pred. No. 5.le-108;
Matches 298; Conservative 77; Mismatches 93; Indels 21; Gaps 8;

Qy 1 QTCPSVCSNQSFKVVCVRKLNREVPDGIISTNRLNLHENQIQIKVNSFKHLRHLLEI 60
Db 44 qncpsvcsnqsfkvvctrrglsevpqgipnsnrylnlmenniqmqaadfrhlhlev 103
Qy 61 LQLSNRHTRTEIGAFNGLANLNTLEFDNRLTTPNGAFVYLSKLELWLRNPNIESIP 120
Db 104 lqgrnsrlqevafnglasintlelfdnwtlvpqafeylskirelwrnnpiesip 163
Qy 121 SYAFNRPLSLRRLDGLKRLSYISEGAFEGLSNRLRYLNLMCNLRTEIPNLTPLIKDEL 180
Db 164 syafnrpslmlrdlgelkklleyisegafeglnlkylnlmgcnkdmnpnlplvgleel 223
Qy 181 DLSGNHLSAIRPQSGFQGLHKLWMTQSQIQVIERNAFDNLQSLVEINLAHNLTLLPH 240
Db 224 emsgnhfpeirpghfghlsslklwvmsqvsliernafdglsalveinlahnlsip 283
Qy 241 DLFTPLHLERLHLHNPWCNCDLTLWSWIKDAPSNATACCARPPLNKGRIYIGEL 300
Db 284 dlftplylvelhlnhpwncddilwawlrveiptnscgchpamhrgylvev 343
Qy 301 DONFTCYAPVIVPEPPADLVTEGMAELKCRSTLSVSWITPNGVMTGHGAYKVRIA 360
Db 344 dqasfqsaptmdaprdlnisegrmaelkr-tppmssvkwllpngtvlshasrhris 402
Qy 361 VLSQGTNFTNTVQDTGMYTCWYSNSVGNVTASATLNVTA--TTTPFSYFSVTVTM 418
Db 403 vldngtlnfshvldstgvtcmvtnvagnsnasaylnvstaeltntsnysftvtvett 462
Qy 419 EPSQDEARTTNNVGPFPVDWETNV-----TTSLTP--QSTRSTKFTTIPVTDINS 470
Db 463 elsped---tirkykvp-----tsttgyqpayttstvtlqtr-vpkqvavpatdtd 513
Qy 471 GI-PGIDEV 478
Db 514 kmqtsldev 522

RESULT 13
AAB24073
ID AAB24073 standard; Protein; 653 AA.
XX
AC AAB24073;
XX
DI 29-JAN-2001 (first entry)

XX Human PRO1111 protein sequence SEQ ID NO:46.
DE Human; tumour; diagnosis; neoplastic disease; neoplastic cell growth;
KW proliferation; tumorigenesis; identification; cancer; cytostatic;
KW neutrotic; neuroprotective; antinflammatory; immunosuppressive;

KW immunostimulant; antiangiogenic; leukaemia; lymphoid malignancy;
KW neuronal disorder; glial disorder; astrocytal disorder; angiogenic;
KW hypothalamic disorder; glandular disorder; macrophagal disorder;
KW epithelial disorder; stromal disorder; blastocoelec disorder;
KW inflammatory disorder; immunologic disorder.
OS Homo sapiens.
XX
PN WO2000053755-A2.
XX
PD 14-SEP-2000.
XX
PF 06-JAN-2000; 2000WO-US00376.
XX
PR 08-MAR-1999; 99WO-US05028.
PR 02-JUN-1999; 99WO-US12252.
PR 23-JUN-1999; 99US-0141037.
PR 07-JUL-1999; 99US-0143048.
PR 26-JUL-1999; 99US-0145698.
PR 30-NOV-1999; 99WO-US28313.
PR 20-DEC-1999; 99WO-US30911.
PR 05-JAN-2000; 2000WO-US00219.
XX

(GETH) GENENTECH INC.

PA Ashkenazi AJ, Baker KP, Goddard A, Gurney AL, Hillan KJ, Roy MA;
PI Watanabe CK, Wood WI;
XX
XX WPI; 2000-572270/53.
DR N-PSDB; AAC58383.

XX Thirty PRO polynucleotides encoding PRO polypeptides, useful in the
PT treatment, diagnosis and prevention of cancer -
PT
XX
PS Claim 61; Fig 34; 286pp; English.

XX The present invention describes an isolated antibody that binds to
CC one of the human PRO proteins designated PRO212, PRO290, PRO341, PRO535,
CC PRO619, PRO717, PRO809, PRO830, PRO848, PRO943, PRO1005, PRO1009,
CC PRO1025, PRO1030, PRO1097, PRO1107, PRO1111, PRO1153, PRO1182, PRO1184,
CC PRO1187, PRO1281, PRO23, PRO39, PRO834, PRO1317, PRO1710, PRO2094, cell
CC PRO2145 OR PRO2198. PRO antagonists can be used to inhibit tumour cell
CC growth. The PRO polypeptides and nucleotides are useful in the
CC treatment, diagnosis and prevention of cancer. The antibodies and other
CC anti-tumour compounds may be used to treat various conditions, including
CC those characterised by overexpression and/or activation of the amplified
CC PRO genes. Exemplary conditions or disorders to be treated with such
CC antibodies and other compounds include benign or malignant tumours
CC (e.g., renal, liver, kidney, bladder, breast, gastric, ovarian,
CC colorectal, prostate, pancreatic, lung, vulva, thyroid, hepatic
CC carcinomas, sarcomas, glioblastomas, and various head and neck tumours),
CC leukaemias and lymphoid malignancies, other disorders such as neuronal,
CC glial, astrocytal, hypothalamic and other glandular, macrophagal,
CC epithelial, stromal and blastocoelec disorders, and inflammatory,
CC angiogenic and immunologic disorders. AAC58242 to AAC58366 represent PCR
CC primers and hybridisation probes used in the isolation of the human PRO
CC sequences. AAC58367 to AAC58396 and AAB24057 to AAB24089 represent human
CC PRO polynucleotide and protein sequences given in the exemplification of
CC the present invention.
XX
SQ Sequence 653 AA;

Query Match 62.7%; Score 1581.5; DB 21; Length 653;
Best Local Similarity 60.9%; Pred. No. 5.le-108;
Matches 298; Conservative 77; Mismatches 93; Indels 21; Gaps 8;

Qy 1 QTCPSVCSNQSFKVVCVRKLNREVPDGIISTNRLNLHENQIQIKVNSFKHLRHLLEI 60
Db 44 qncpsvcsnqsfkvvctrrglsevpqgipnsnrylnlmenniqmqaadfrhlhlev 103
Qy 61 LQLSNRHTRTEIGAFNGLANLNTLEFDNRLTTPNGAFVYLSKLELWLRNPNIESIP 120
DB 104 lqgrnsrlqevafnglasintlelfdnwtlvpqafeylskirelwrnnpiesip 163

PR 04-AUG-1998; 98US-0095301.
PR 04-AUG-1998; 98US-0095302.
PR 04-AUG-1998; 98US-0095318.
PR 04-AUG-1998; 98US-0095321.
PR 04-AUG-1998; 98US-0095325.
PR 10-AUG-1998; 98US-0095916.
PR 10-AUG-1998; 98US-0095929.
PR 10-AUG-1998; 98US-0096012.
PR 11-AUG-1998; 98US-0096143.
PR 11-AUG-1998; 98US-0096146.
PR 12-AUG-1998; 98US-0096329.
PR 12-AUG-1998; 98US-0096757.
PR 17-AUG-1998; 98US-0096766.
PR 17-AUG-1998; 98US-0096768.
PR 17-AUG-1998; 98US-0096773.
PR 17-AUG-1998; 98US-0096791.
PR 17-AUG-1998; 98US-0096867.
PR 17-AUG-1998; 98US-0096891.
PR 17-AUG-1998; 98US-0096894.
PR 17-AUG-1998; 98US-0096895.
PR 17-AUG-1998; 98US-0096897.
PR 18-AUG-1998; 98US-0096949.
PR 18-AUG-1998; 98US-0096950.
PR 18-AUG-1998; 98US-0096959.
PR 18-AUG-1998; 98US-0096960.
PR 18-AUG-1998; 98US-0097022.
PR 19-AUG-1998; 98US-0097141.
PR 20-AUG-1998; 98US-0097218.
PR 20-AUG-1998; 98US-0097261.
PR 26-AUG-1998; 98US-0097951.
PR 26-AUG-1998; 98US-0097952.
PR 26-AUG-1998; 98US-0097954.
PR 26-AUG-1998; 98US-0097955.
PR 26-AUG-1998; 98US-0097971.
PR 26-AUG-1998; 98US-0097974.
PR 26-AUG-1998; 98US-0097978.
PR 26-AUG-1998; 98US-0097979.
PR 26-AUG-1998; 98US-0097986.
PR 26-AUG-1998; 98US-0098014.
PR 31-AUG-1998; 98US-0098525.
PR 16-SEP-1998; 98US-0100634.
PR 12-JAN-1999; 99US-0115565.
XX
XX
XX (GETH) GENENTECH INC.
XX Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK;
XX Wood WJ, Yuan J;
XX
XX WPI; 2000-072883/06.
XX N-PSDB; AAZ65033.
XX
XX Membrane-bound proteins and related nucleotide sequences
XX
XX claim 12; Fig 157; 822pp; English.
XX
XX The invention provides membrane-bound PRO polypeptides and
XX polynucleotides encoding them. The PRO sequences of the invention were
XX identified based on extracellular domain homology screening. The PRO
XX sequences have homology with proteins including LDL receptors, TIE
XX ligands and various enzymes. The membrane-bound proteins and receptor
XX molecules are useful as pharmaceutical and diagnostic agents. Receptor
XX immunoadhesins, for instance, can be used as therapeutic agents to block
XX receptor-ligand interactions. The membrane-bound proteins can also be
XX employed for screening of potential peptide or small molecule inhibitors
XX of the relevant receptor/ligand interaction. The PRO encoding sequences
XX are useful as hybridization probes, in chromosome and gene mapping and in
XX the generation of antisense RNA and DNA. PRO nucleic acid sequences
XX will also be useful for the preparation of PRO polypeptides, especially
XX by recombinant techniques.
XX
XX Sequence 653 AA;

Query Match 62.7%; Score 1581.5; DB 21; Length 653;
Best Local Similarity 60.9%; Pred. No. 5.1e-108;
Matches 298; Conservative 77; Mismatches 93; Indels 21; Gaps 8;
Qy 1 QTCPSVSCSNQFSKVICVRKNLREVPDGI STNTRLLMLHENQIOIIVKNSFKHRLHLEI 60
Db 44 qncpsvscsnqfskvctrrglsevpqgipnsntrlylmennlqmqadttrfhlhlev 103
Qy 61 LQLSRHHRTTETGAFNGLANLTLELFDNRLTTPNGAFVYLSKLKELWRNNPIESTP 120
Db 104 lqlgrnsirqlievafnglasintlielfdnwltvlpagafeylsklrelwlrnnpieslp 163
Qy 121 SYAFNRIPSLRRDLGELKRLSYISEGAFEGLSNLRYLNLAMCNLRLEIPNLPLIKLDEL 180
Db 164 syafnrpslrmldlgelkklkleyisegafeglnfkylnlmgcnlikdnpnltpvgleel 223
Qy 181 DLSGNHLSAIRPGSFQGLMHLOKLMWIOSQIOVIERNAPDNLSQSLVEINLAHNNTLLPH 240
Db 224 emsgnhfpeirpgrsfhglsslkklwmnsqvsliernafdglsalveinlahnnlssiph 283
Qy 241 DLFTHPLHLERHLHHPNPNCCDILWLSWIKDMAPSNATACCARCNTPPNLLKGRYIGEL 300
Db 284 dlftprlyivelhlnhnpwncddilwawlrreypcnstccgrchpmmrgrylvev 343
Qy 301 DQNYFTCYAPVIVEPPADLVNTEGMAELKCRASTSLTSVSWITPMTGVMTHGAYKVRIA 360
Db 344 dqasfqcscapfimdaprdlnisegmaelkcr-tpmssvkwllpngctvshashrpris 402
Qy 361 VLSGDTLNFTNVTVDTCMYTCWNSVGNVTASATLNVTA--TTTFPSFSTVVTETM 418
Db 403 vlngdtlnfshvllsdgtvytcmtvnavgnasnasylnvstaelnsgysffttvtvtt 462
Qy 419 EPSODEARTTNNVGPVPVDWETTNV-----TTSLP--QSTRSTEKTFPTIDINS 470
Db 463 eisped---trkykvpv-----ttstgyqpayttsttvlqtr-vpkqvavpatdtd 513
Qy 471 GI-PEIDEV 478
Db 514 kmqtsldev 522
RESULT 15
AAE09438
ID AAE09438 standard; Protein; 653 AA.
XX AC AAE09438;
XX
XX 19-NOV-2001 (first entry)
XX
XX Human sbgPRO331a protein.
XX
XX Human; Alzheimer's disease; amyotrophic lateral sclerosis;
XX ALS; Zollinger-Ellison syndrome; immune system disease; schizophrenia;
XX inflammation; hematopoietic disease; anxiety; feeding disorder; aging;
XX anorexia; depression; cardiovascular disease; sleep disorder; seizure;
XX memory alteration; migraine; stroke; asthma; neuropathy; hypoglycaemia;
XX sexual disorder; growth abnormality; infection; autoimmune disease;
XX rheumatoid arthritis; cataractogenesis; angiogenesis; atherosclerosis;
XX cerebral ischaemia; cirrhosis; Huntington's disease; Hodgson's disease;
XX hypercholesterolaemia; headache; amnesia; cardiac arrhythmia; obesity;
XX diabetes mellitus; glomerulonephritis; renovascular hypertension;
XX cancer; vaccine; gene therapy; sbgPRO331a gene.
XX
XX Homo sapiens.
XX
XX WO200160850-A1.
XX
XX 23-AUG-2001.
XX
XX 14-FEB-2001; 2001WO-US04703.
XX
XX 14-FEB-2000; 2000US-0182172.
XX 29-FEB-2000; 2000US-0186084.
PR

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 15:44:38 ; Search time 20.91 seconds
(without alignments)

US-09-905-056-292_COPY_45_522
2524
Title: Perfect score:
Sequence: 1 OTCPSCVCSNOFSKVICVR.....KTFTIPVTDINSIGPGDEV 478

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```
Database :      PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	1151.5	45.6	421	2	T46266	hypothetical prote
2	402.5	15.9	1091	2	A58532	glial cell membran
3	336	15.7	707	2	JC7763	neuronal leucine-r
4	394	15.6	361	2	A53860	chondroadherin pre
5	389.5	15.4	1531	2	T42218	slit-1 protein hom
6	365.5	14.5	1523	2	T13953	MEGF5 protein - ra
7	350	13.9	1469	2	B36665	slit protein 2 pre
8	350	13.9	1480	2	A36665	slit protein 1 pre
9	345	13.7	789	2	T28714	hypothetical prote
10	345	13.7	1355	2	T28715	hypothetical prote
11	344.5	13.6	1025	2	T42626	secreted leucine-r
12	329.5	13.1	560	2	A60164	platelet membrane
13	329.5	13.1	605	2	JC5239	insulin-like growt
14	314.5	12.5	605	2	A41915	insulin-like growt
15	313	12.4	1535	2	S46224	peroxidase - frui
16	309	12.2	603	2	JC1282	insulin-like growt
17	302.5	12.0	603	2	JG6128	insulin-like growt
18	302	12.0	359	1	NBHUC8	decorin precursor
19	293	11.6	360	2	S06280	decorin precursor
20	291	11.5	357	2	S24317	decorin precursor
21	290.5	11.5	594	2	T23841	hypothetical prote
22	289	11.5	360	2	I47020	decorin - rabbit
23	286.5	11.4	536	2	A34901	lysine carboxypept
24	286	11.3	354	2	A55454	decorin precursor
25	283.5	11.2	420	2	A53531	oncofetal tropobla
26	283	11.2	354	2	S23145	decorin precursor
27	282	11.2	626	1	NBHUIA	platelet glycoprot
28	281.5	11.2	610	2	T23836	hypothetical prote
29	281.5	11.2	1389	2	T13852	genewheeler prote

ALIGNMENTS

RESULT

T46266
KLS011

140200
hypothetical protein DKFZp761A179.1 ~ human (fragment)

C:Species: Homo sapiens (man)

C, Species: homo sapiens (man)
C: Date: 04-Feb-2000 #sequence revision 04-Feb-2000 #text change 04-Feb-2000

C: Accession: T46266

C, Accession: 140200
R: Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

K/Blum, H.; Baderisachs, S.; Newes, H.W.; Baderisachs, S. submitted to the Protein Sequence Database, January 2000

A:Reference number: Z23034

A; Accession: T46266

A:Status: preliminary

A; Molecule type: mRNA

A; Residues: 1-421 <AAA>

A;Cross-references: EMBL:AL137451

A; Experiment

C;Genetics:

Query Match	45.6%;	Score 1151.5;	DB 2;	Length 421;
Best Local Similarity	59.8%;	Pred. No. 3.3e-72;		
Matches 213:	Conservative	53: Mismatches	51:	Indels 39: Gaps 6:

Qy	162	MCNLRPNUTPLIKKDELDSGNHLSA	IRPGSFQGLMHLOKLUWMTQSOIQVIERNAFD	221
Db	1	MCNKDIPNETALVRLEELS	NRDLIRPGSFQGLTSURKLMLHAQVATIERNAFD	60
Qy	222	LOSIVETNAHNNLTLLPHDLFTPLHLEIR	HLUHHNPWNCNCIDILWSWIKDMAPSNTA	281
Db	61	LKSLEELNLSHNNLMSLPHDLFTPLHLRLER	VHLNHPWCHNCDDVLMLSWKLKETSPTSNT	120
Qy	282	CCARCNTPPNLKGRYIGELDONFTCYAPV	IVPEPPADLVNTEGMAAELKCRASSTLSUTSVS	341
Db	121	CCARCHAPAGLKGRYIGELDQSHFTCYAPV	IVPEPPTDLNVTEGMAAELKCRTGSTMTSVN	180
Qy	342	WTPNGVMTPHGAKYRIAVLSDGTLNFTNV	TQDTCWYTCWVSNVSGNTASATLNVTA	401
Db	181	WLTPNGTLMTHGSTRVRSVLHBDGTLNFT	NVTQDTCQYTCWVNSAGNTASATLNVSA	240
Qy	402	ATTPP-----	FSYFTSVTIVTME--PSQD--EARTDNN-VG	433
Db	241	VDPAAGTGGGGGCGSGVGGGGGGYFTYFT	VTVTVETLLETPGGEALOPRGTEKEPPG	300

RESULT

A58532

A38532
glial cell membrane glycoprotein [JG-1 precursor - mouse

C: Species: *Mus musculus* (house mouse)

C;Date: 11-Apr-1997 #sequence_revision 11-Apr-1997 #text_change 05-Nov-1999
C;Accession: A58532

R;Suzuki, Y.; Sato, N.; Tohyama, M.; Wanaka, A.; Takagi, T.

J. Biol. Chem. 271, 22522-22527, 1996

A;Title: cDNA cloning of a novel membrane glycoprotein that is expressed specifically in

A;Reference number: A58532; MUTID:96394313

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-1091 <SUZ>

A;Cross-references: GB:D78572; NID:g1545806; PIDN:BAAL1416.1; PID:g1545807

C;Superfamily: leucine-rich alpha-2-glycoprotein repeat homology; proteoglycan amino-ter

F;36-61/Domain: leucine-rich alpha-2-glycoprotein repeat homology <PAH>

F;95-117/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>

F;118-141/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>

F;142-165/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>

F;166-189/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>

F;191-213/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>

F;214-237/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>

F;238-261/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>

F;262-285/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>

F;286-309/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>

F;310-333/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>

F;334-357/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR11>

F;358-381/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR12>

F;385-408/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR13>

F;409-432/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR14>

F;440-485/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR15>

F;440-485/Domain: proteoglycan carboxyl-terminal homology <PCH>

Query Match 15.9%; Score 402.5; DB 2; Length 1091;

Best Local Similarity 24.1%; Pred. No. 7.5e-20;

Matches 134; Conservative 77; Mismatches 168; Indels 177; Gaps 14;

QY 37 LNLHENQIITKVSFKLRLHLEILQISRNHRIETGAFNGLANLTLLEFDNRLTIP 36

Db 218 LDLRNRRLLEGUTFOGLDSLEVLRLQRNNISRLTDGAFGLSKMHLVHLEYSLEVN 277

QY 97 NGAFVYLSKELMLRLNPIESIPS-----YAPNRIPIPSRLRLDLGELKRL-- 141

Db 278 SGSLYGLTALHQLHLSNNSISRIQDGSFCQKLHLEILSFNNLRLDEESLSSLSI 337

QY 142 -----SYISEGAFGLSNRLYNLAMCNLRPEINLTPLIKDELDLSNHLISATRP 193

Db 338 LRLSHNAISHIAEAGFKLSRLVD-----LDHNEISGTIEDT--SG 378

QY 194 SFOGLMHLQKLMWTOSQIOVTERNAFONLSLVEINLAHNNLTLLPHDLFTPLHLERIH 253

Db 379 AFTGLDNLKSLTLFGNKIKSVAKRAFSGLEHLNLGSENAIRSVQDFAFKMLNKELY 438

QY 254 LHHNPWNCNDILMSLWMI--KDMAPSNATACCRCNTPPNPKRGYI--GELDONY----- 304

Db 439 ISSSEFLCDCLQKWLPPWLMGRMLQAFVATCAH--PESLKGOSIFSLVPDSFVCDFF 495

QY 305 -----FTCYA----- 309

Db 496 KPQIITPETTMVAVGKDIFRTCSAASSSSPMTFANKKNEVLANADMFHVAHQDG 555

QY 310 -----PVIVEPPADLNVTG 324

Db 556 EVMEYTTILHURHVTFGHEGRYQCIINHFSGTYSYSHKARLTVNVLPSFTKIPHDIAIRTG 615

QY 325 MAELKCRASLSLT--SVSWITPNTGVTMTHGAYKVRIAVLSLDTLNF--TNVTVDQTMVTC 382

Db 616 TTABLECAATGHPNPQATQWQDGTDP--AARERMHVMPDGDVFFITDVKIDDMGVYSC 674

QY 383 MYSVSGNVTASATLVNTAATTFPSYFSTVVTETMPSQDEARTDNV-----GP 434

Db 675 TAQNSAGSVSANATLTLE-----TPSLAVPLEDRVTVTGTVAFOCKATGSP 722

QY 435 TPVVDWETNTTTSIT 450

Db 723 TPRITWLKGGRPRLSLT 738

RESULT 3

JC7763

neuronal leucine-rich repeat protein-3 - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002

C;Accession: JC7763

R;Fukanachi, K.; Matsuo, Y.; Kitanaka, C.; Kuchino, Y.; Tsuda, H.

Biochem. Biophys. Res. Commun. 287, 257-263, 2001

A;Title: Rat neuronal leucine-rich repeat protein-3: Cloning and regulation of the ge

A;Reference number: JC7763; PMID:11549284

A;Contents: Fibrosarcoma cells

A;Accession: JC7763

A;Molecule type: mRNA

A;Residues: 1-707 <FUK>

A;Cross-references: GB:AF291437

C;Comment: This protein, a new member of the neuronal leucine-rich repeat protein fam

in protein-protein interaction and functions as a cell adhesion molecule or soluble l

C;Genetics:

A;Gene: nlrr-3

C;Keywords: cell adhesion

Query Match 15.7%; Score 396; DB 2; Length 707;

Best Local Similarity 25.9%; Pred. No. 1.2e-19;

Matches 127; Conservative 79; Mismatches 166; Indels 98; Gaps 13;

QY 3 CPVSCSCS-----NOFSKVICVRKNLREVPGDGIISNTNRLNHNHQIILK--- 48

Db 29 CPQLCTCEIRPWFPTPSIYMEATVDCNDGLNFPARLPADTQILLQTNNIARIHST 88

QY 49 -----VNSFKHLRHLLEILQISR-----NHIRTIEIGAFNGLANLTLELF 88

Db 89 DFPVNLGTGLDSQNNLSVTNINQKSSQLLSVLEENKLTPEKCLYGLSNLQELIYN 148

QY 89 DNRUTTPNGAFVYLSKELMLRLNPIESIPSYAFNRIPIPSRLRLDLGELKRLSYISEGA 148

Db 149 HNLISATSPGAFVGLHLLRLHLNLSNRLQMNKWFALPNLEILMLGD--NPILRIKDMN 207

QY 149 FEGLSNRLYNLAMCNLRPEI-----NLTPLIKDE-----LDL 182

Db 208 FQPLKRLSLVIAGINTETVDPDLDALVLENLESTFYDNRLNKPQVALQKAVNLKFLDL 267

QY 183 SGNHLSATSPGSGFQGLMHLQKLMW-----IQS-----QIQVTER 216

Db 268 NKNINRIIRRGDFSNMHLKELGINNMPPELVSDSLAVDNLPLDKRIEATNPRLSYIHP 327

QY 217 NAFNLQSLVEINLAHNNLTLLPHDLFTPLHLHLRIHLHNPWNCNDILMSW---WIK 273

Db 328 NAFERLPKLESIMLSNLSALSALYHGTIESLPNLKEISHSNPIRCDVIRWINMKNKTNR 387

QY 274 DMAFSNTACCRCNTPPNPKRGYIGELD--QNYFTCYAPVIVEP--PADLNVTGMAAEK 330

Db 388 FMEPDSLFCV---DPPEFQGNVRQVFRDMTEICLPDIAPESFPSTLDEADSVSLH 443

QY 331 CRASTG--LTSVSWITPNTGVTMTHGAYKVRIAVLSLDTLNF--TNVTVDQTMVSNVSG 389

Db 444 CRATAEPOPEIYWIPTSGKRLLPNTLREKFYVHSEGLDIRGITPKEGGLTYCIATNLVG 503

QY 390 NTTASATLVN 399

Db 504 ADLKSIMIKV 513

RESULT 4

A53860

chondroadherin precursor - bovine

N;Alternate names: 38K leucine-rich protein

C;Species: Bos primigenius taurus (cattle)

C;Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 05-Nov-1999

C;Accession: A53860

Db 271 PACNANSLSCPSACSCSNNIVDCRGKGLTEIPANLPEGIVEIRLEQNSIKSIPAGAFIOY 330
QY 176 -KLDELDSGNHLSAIRGSFQ-----GLMHQLKLMWIOSQ 210
Db 331 KKLKRIDISKNOISDIADAFQGLKLSLSVLYGNKITEIPKGFDFGLVLSQLLLLNANK 390
QY 211 IQVIERNAFDNLQSLVEINLAHNNLTLLPHDLFTPLHLHLRIHNNPWNCCDILWLSW 270
Db 391 INCLRVNTFQDILQNLNLSLSYDNKLTQTSKGLFAPLQSIQTLHLAQNFVCDCHLKLWAD 450
QY 271 WIKDMAPNTACCARNTPPNLKGRYIGELDONFTFC 307
Db 451 YLQD-NPIETS-GARCSSPRRLANKRISQIKSKFRFC 485

RESULT 7

B36665
slit protein 2 precursor - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 30-Apr-1991 #sequence_revision 30-Apr-1991 #text_change 19-May-2000
C:Accession: B36665
R:Rothberg, J.M.; Jacobs, J.R.; Goodman, C.S.; Artavanis-Tsakonas, S.
Genes Dev. 4, 2169-2187, 1990
A:Title: slit: an extracellular protein necessary for development of midline glia and cc
A:Reference number: A36665; MUID:91099665
A:Accession: B36665
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1469 <ROT>
A:Cross-references: GB:X53959
C:Genetics:
A:Gene: FlyBase:sli
A:Cross-references: FlyBase:FBgn0003425
C:Superfamily: unassigned EGF-related proteins; EGF homology; leucine-rich alpha-2-glyco-
F:66-91/Domain: proteoglycan amino-terminal homology <PAH1>
F:101-124/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F:125-148/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F:149-172/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F:173-196/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F:197-220/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F:228-272/Domain: proteoglycan carboxyl-terminal homology <PCS1>
F:288-313/Domain: proteoglycan amino-terminal homology <PAH2>
F:323-346/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F:347-370/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
F:371-394/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>
F:395-418/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>
F:419-442/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>
F:450-494/Domain: proteoglycan carboxyl-terminal homology <PCS2>
F:512-537/Domain: proteoglycan amino-terminal homology <PAH3>
F:547-571/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR11>
F:572-595/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR12>
F:596-619/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR13>
F:620-643/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR14>
F:651-695/Domain: proteoglycan carboxyl-terminal homology <PCS3>
F:708-733/Domain: proteoglycan amino-terminal homology <PAH4>
F:743-766/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR15>
F:767-790/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR16>
F:846-890/Domain: proteoglycan carboxyl-terminal homology <PCS4>
F:1028-1061/Domain: EGF homology <EGF>
F:1068-1099/Domain: EGF homology <EGF>
F:1115-1148/Domain: EGF homology <EGF1>

Query Match 13.98; Score 350; DB 2; Length 1469;
Best Local Similarity 23.98; Pred. No. 4.8e-16;
Matches 109; Conservative 65; Mismatches 152; Indels 130; Gaps 9;

QY 2 TCPVSCSNFQSVICVRKRLNPEVDPDGISTNRLNLNHNQIQIKVNSFKHLRHLRIL 61
Db 294 SCPHPCRCADGI--VDCREKSLTSVPVLPDDTTDLRLEQNFITLPPKPSFSSFRRLRI 351
QY 62 QLSRHHITIEIGAPNGLANLTLELFDNRLTTPNGAFVYLSKLKELWLRNNPISIPS 121

Db 352 DLSNNNISRIADHDAISGLKQLTTLVLYGNKIKDLPVSGVFKGLSRLLLNANETSCIRK 411
QY 122 YAFNRISRLDLGELKRLSYISEGAPEGLSNRLYNLAM-----CNLR----- 166
Db 412 DAFRDLHSLLSLYD-NNIQLSLANGTFDAMKSMKTVHLAKNPFITCDNLRWLADYLHN 470
QY 167 ----- 166
Db 471 PIETSGARCESPKRMHRRRIEESLREEKFKCSWGLRMKLSGECRMDSDCPAMCHCEGTV 530
QY 167 -----EIPNLTP-----IKLD-----ELDSLGNHLSAIRGS 194
Db 531 DCTGRLKEIPRIDPLHTTELLANDNELGRISSDGLFGRPLHLVLEKLRNQLTGIEPNA 590
QY 195 FOGLMHQLKLMWIOSQIOVIERNAFDNLQSLVEINLAHNNLTLLPHDLFTPLHLHLRTHL 254
Db 591 FEGASHIOELQELQENKIKESNKMPFLGHLQTLNLQNSCVMPGSFEHLNLSLNL 650
QY 255 HNNPWNCCDILWLSWIKDMAPNTACCARNTPPNLKGRYIGELDONFTFCYAPVIVE 314
Db 651 ASNPENCNCHLAWFAECVRYKSLNGGA--ARCGAPSKYRDVQIKDLPHSEFKCS----- 703
QY 315 PPADLVNTEGMAELKCRASSTLSVSWITNGTVM 350
Db 704 -----ENSEGCLGDGYCPPSCTCT-----GTWV 726
RESULT 8
A36665
slit protein 1 precursor - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 30-Apr-1991 #sequence_revision 30-Apr-1991 #text_change 17-Nov-2000
C:Accession: A36665; A31640; S13523
R:Rothberg, J.M.; Jacobs, J.R.; Goodman, C.S.; Artavanis-Tsakonas, S.
Genes Dev. 4, 2169-2187, 1990
A:Title: slit: an extracellular protein necessary for development of midline glia and
A:Reference number: A36665; MUID:91099665
A:Accession: A36665
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1480 <ROT>
A:Cross-references: GB:X53959; NID:g8614; PIDN:CAA37910.1; PID:g8615
R:Rothberg, J.M.; Hartley, D.A.; Walther, Z.; Artavanis-Tsakonas, S.
Cell 55, 1047-1059, 1988
A:Title: slit: An EGF-homologous locus of D. melanogaster involved in the development
A:Reference number: A31640; MUID:89077533
A:Accession: A31640
A:Molecule type: DNA
A:Residues: 881-1182, 'G', 1185-1404, 'GT', 1463-1464, 'YHA' <RO2>
A:Cross-references: GB:M23543; NID:g340939; PID:g514357
C:Genetics:
A:Gene: FlyBase:sli
A:Cross-references: FlyBase:FBgn0003425
A:Introns: 1351/3
C:Superfamily: unassigned EGF-related proteins; EGF homology; leucine-rich alpha-2-gly-
C:Keywords: alternative splicing; growth factor
F:66-91/Domain: proteoglycan amino-terminal homology <PAH1>
F:101-124/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F:125-148/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F:149-172/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F:173-196/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F:197-220/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F:228-272/Domain: proteoglycan carboxyl-terminal homology <PCS1>
F:288-313/Domain: proteoglycan amino-terminal homology <PAH2>
F:323-346/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F:347-370/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
F:371-394/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>
F:395-418/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>
F:419-442/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>
F:450-494/Domain: proteoglycan carboxyl-terminal homology <PCS2>
F:512-537/Domain: proteoglycan amino-terminal homology <PAH3>
F:547-571/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR11>

F:572-595/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR12>
F:596-619/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR13>
F:620-643/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR14>
F:651-695/Domain: proteoglycan carboxyl-terminal homology <PCS3>
F:708-733/Domain: proteoglycan amino-terminal homology <PAH4>
F:743-766/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR15>
F:767-790/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR16>
F:791-814/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR17>
F:815-838/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR18>
F:846-890/Domain: proteoglycan carboxyl-terminal homology <PCS4>
F:1028-1061/Domain: EGF homology <EGF>
F:1068-1099/Domain: EGF homology <EGF2>
F:1115-1148/Domain: EGF homology <EGF1>

Db 280 SNTLAVCVEDGAVLYNTSMPLRLSRFTNNQLRYPKRAFEPFALEELDTDNPIATIH 339
Qy 240 HDLFTPLHLHRIHLHNPWCNCOIWLWIIKMDAPSNTPACCARNTPPNLRKGRIGE 299
Db 340 PEAFEPPL-ELKRLVMNSSIIICDCQISLWASWIIYRLDKKSIIAKGSYPPPLADLYVVA 398
Qy 300 LDQNVFTCY----APVIVEPPADLVNTEGMAAELKCRA-STSLTSVSW----- 342
Db 399 IDTANLCHNDSPRAKIVRQVPEVSTLIGEKARTCNVYGASPLSIEMRWENQOPRVLV 458
Qy 343 -----TPPNGTVMTHGAYKRYAIVASDGLTNFTNVVQDGTGMTVCMVNSVGN 390
Db 459 QDSATFELSINRTAVVNGTFDER--ELAAAEELLDNVAMTDNSEYQCVRNRFGS 510

RESULT 11
T42626
secreted leucine-rich repeat-containing protein SLIT2 - mouse (fragment)
N:Alternate names: neurogenic extracellular slit protein
C:Species: Mus musculus (house mouse)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-Jan-2000
C:Accession: T42626
R:Holmes, G.P.; Negus, K.; Burridge, L.; Raman, S.; Algar, E.; Yamada, T.; Little, M.H.
Mech. Dev. 79, 57-72, 1998
A:Title: Distinct but overlapping expression patterns of two vertebrate slit homologs in
A:Reference number: 222177; MUID:99279238
A:Accession: T42626
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1025 <HOL>
A:Cross-references: EMBL:AF074960; NID:g4151258; PID:g4151259; PIDN:AAD04345.1
C:Genetics:
A:Gene: Slit2

Query Match 13.6%; Score 344.5; DB 2; Length 1025;
Best Local Similarity 24.3%; Pred. No. 7.1e-16;
Matches 102; Conservative 62; Mismatches 149; Indels 107; Gaps 8;

Qy 3 CPVSCVCSNQFSKVICYKRNLRVPGDGISTNTRLLNLHENQIOLIKVNS-FKHLRHLEIL 61
Db 2 CPEKCRCEG--TTVDCSNORLNKIPDHIPQVTAELRNNEFTVLEATGIFKLPQLRXI 59
Qy 62 QLSNRHIRTIBIGAFNGLANLTLELFDNRLTTPNGAFVYLSKLKELWLRNPNIESIPS 121
Db 60 NFSNKKITDIEGAFEGAGSGVNEILLTSNRLNENYQHKFKGLESKTLMLRSNRISCVGN 119
Qy 122 YAFNRPSLRKDLGELKRLSYISEGAFEGLSNLRNYNLAM-----CN----- 164
Db 120 DSFTGLGSVRLSLYD-NQITTVAPGAFDXLHLSLSTLNLANPNCNCHLAWLGELRRK 178
Qy 165 -----LREIP-----NLTPLIK----- 176
Db 179 RIVTGNPRCKPYELKEIPIQDVAIQDFTCDGNDNDCSPLSRCPSECLDFTYVRCSN 238
Qy 177 -----LDELDSGNHLSAIRPGSFQGLMH 200
Db 239 KGLKVLPGIKPVTELYLDGNQFTLPKELSNYKHLTLIDLNSNRISTLSNQXFSNWTQ 298
Qy 201 LQKLMWTQSQIQVTERNAFNQLSLVEINLAHNNLTLLPHDLFTPLHLHLHHHPWN 260
Db 299 LTLILSYNRLRCIPPTFDGLSKRLSLHSGNDISVVPECAFNDLSALSHLAIGANPLY 348
Qy 261 CNCIDILMSWIKDMASNTACCARCNTPPNLRKRYIGELDONVFTCYAPVIVEPPADLN 320
Db 359 CDCNQWLSDDWK--SEYKEPGTARCAGPGEMADKLLLTTPSKKFTCGGPMDTIIQACN 416

RESULT 12
A60164
platelet membrane glycoprotein v precursor - human
C:Species: Homo sapiens (man)

C:Date: 12-Jan-1993 #sequence_revision 24-Feb-1994 #text_change 05-Nov-1999
C:Accession: A48030; A60164; A35483; B35483; C35483; A60432; A47507; S34329
R:Lanza, E.; Morales, M.; de la Salle, C.; Cazenave, J.P.; Clemenson, K.J.; Shimomura
J. Biol. Chem. 268, 20801-20807, 1993
A:Title: Cloning and characterization of the gene encoding the human platelet glycoprotein
A:Reference number: A48030; MUID:94012616
A:Accession: A48030
A:Molecule type: DNA
A:Residues: 1-560 <LA2>
A:Cross-references: EMBL:223091; NID:g312501; PIDN:CAA80637.1; PID:g312502
R:Shimomura, T.; Fujimura, K.; Maehama, S.; Takemoto, M.; Oda, K.; Fujimoto, T.; Oyama
Blood 75, 2349-2356, 1990
A:Title: Rapid purification and characterization of human platelet glycoprotein V: th
A:Reference number: A60164; MUID:90275263
A:Accession: A60164
A:Molecule type: protein
A:Residues: 365-384, 'X', 386-390, 'X', 392-395, 'X', 397, 188-208, 'I', 210, 27-50, 'X', 52-53, 1
, 'XX', 108, 'T', 161-72, 'TK', 75-77, 'V', 56-57, 'G', 479-487, 'X', 489-496, 'X', 500, 'X', 502-503,
R:Roth, G.J.; Church, T.A.; McMullen, B.A.; Williams, S.A.
Biochem. Biophys. Res. Commun. 170, 153-161, 1990
A:Title: Human platelet glycoprotein V: a surface leucine-rich glycoprotein related t
A:Reference number: A35483; MUID:90321220
A:Accession: A35483
A:Molecule type: protein
A:Residues: 145-166, 'I', 168-169, 'X', 171-172 <ROT>
A:Note: this proteolytic fragment was designated peptide M392
A:Accession: B35483
A:Molecule type: protein
A:Residues: 121-129, 'W', 131-135, 466-468, 'X', 470 <RO2>
A:Note: this material was designated peptide M393 but may contain two peptides
A:Accession: C35483
A:Molecule type: protein
A:Residues: 252-266, 'H', 268-272, 'X', 274-279, 'I', 281-284, 'I', 286 <RO3>
A:Note: this proteolytic fragment was designated peptide M401
R:Zafar, R.S.; Walz, D.A.
Thromb. Res. 53, 31-44, 1989
A:Title: Platelet membrane glycoprotein V: characterization of the thrombin-sensitive
A:Reference number: A60432; MUID:89162331
A:Accession: A60432
A:Molecule type: protein
A:Residues: 477-478, 'FX', 481-485, 'E', 487, 'V', 489-492, 'NQ', 495, 'E', 497-498 <ZAF>
R:Hickey, M.J.; Hagen, F.S.; Yagci, M.; Roth, G.J.
Proc. Natl. Acad. Sci. U.S.A. 90, 8327-8331, 1993
A:Title: Human platelet glycoprotein V: characterization of the polypeptide and the r
A:Reference number: A47507; MUID:93391348
A:Accession: A47507
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-560 <RES>
A:Cross-references: GDB:11238; NID:g388759; PIDN:AAA03069.1; PID:g388760
C:Comment: This platelet membrane protein is a substrate for thrombin.
C:Comment: The amino end of the intact protein is blocked.
C:Comment: This protein is absent in Bernard-Soulier syndrome.
C:Genetics:
A:Gene: GDB:GP5
A:Cross-references: GDB:230236; OMIM:173511
A:Map position: 5pter-5qter
C:Superfamily: leucine-rich alpha-2-glycoprotein repeat homology
C:Keywords: blocked amino end; glycoprotein; platelet; tandem repeat; transmembrane p

Query Match 13.1%; Score 329.5; DB 2; Length 560;
Best Local Similarity 29.7%; Pred. No. 3.5e-15;
Matches 83; Conservative 43; Mismatches 104; Indels 49; Gaps 3;

Qy 22 NLREVDPG-----ISTNTRLLN-----LHENQIOLIKVNSFKH 54
Db 181 NLTHPLGGLLGAQAKLERLLHNSRLSLDGLNSLGTALTELFQFHRNHIRSTAPGADR 240
Qy 55 LRHLEILQLSRNHRITTEIGAFNGLANLTLELFDNRLTTPNGAFVYLSKLKELWLRN 114
Db 241 LPNLSSTLSRNLHAFPLPSALFLHSHNLTLTLFENPLAELPGVLFGEMLGQLDELNRT 300

R;Nelson, R.E.; Fessler, L.I.; Takagi, Y.; Blumberg, B.; Keene, D.R.; Olson, P.F.; Parke
EMBO J. 13, 3438-3447, 1994
A:Title: Peroxidase; a novel enzyme-matrix protein of Drosophila development.
A:Reference number: S46224; MUID:94341255
A:Accession: S46224
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1535 <NEL>
A:Cross-references: GB:U11052; NID:g531384; PIDN:AAA61568.1; PID:g531385
C:Superfamily: peroxidase; myeloperoxidase homology; proteoglycan amino-terminal homolo
F:19-44/Domain: proteoglycan amino-terminal homology <PAH4>
F:661-1350/Domain: myeloperoxidase homology <MPX>

```
Query Match          12.4%   Score 313; DB 2; Length 1535;
Best Local Similarity 24.7%   Pred. No. 1.8e-13;
Matches 97; Conservative 52; Mismatches 136; Indels 108; Gaps 12;

QY  3 CPVSCGCSNOFSKVCYKRLREVPDGIISTNRLNLHENOIQIKVNSFKHLRHLILQ 62
   ||: ||: : ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db  26 CPAGCTCLER--TVRCIRAKLSAVP-----KLPQDTQTLD 58

QY  63 LSRNHITIEIGANGLANLTLEFNRLTTPNGAFVYLSKLELWLNPNIESIPSY 122
   | ||: : ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db  59 LRFNHIEELPANAFSGLAQLTTLFLNDNELAYLQDGLNGLTALRFVYLANNNRLSLPAT 118

QY  123 AFNRIPSLRRLDLGELKRLSVISEGAPEGLSNRLRYLNAMCNLREIPNLTKLDEL 182
   | ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db  119 IFQRMPL-----EGIF--LEN----- 133

QY  183 SGNHLSAIRPGSFGMLHLOKLMWIOQIOVIERNAFDNLQSLVEINLAHNNLTLLPHDL 242
   ||: ||: : ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db  134 -----NDIMOLPAGL-----FDNLPRLNRLIMYNNKLTQLPVDG 167

QY  243 FTPLHLHLRIHLHN--PWNCNDI--LWLSWIKDMAPSNACCACRCNTPPNLKGRYIG 298
   | ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db  168 FNRLNNLRLRLDGNADIDCNCGVYSLWRWHLL-DVQROLVSLTCAAPQMLQNGFS 226

QY  299 ELQNYFTCYAPVIVEPPADINVTGMAAEKLCRASTSL--TSVSWITPNGTVMTHGAYK 356
   | ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db  227 SLGEHFKCAKPKQLVAPQDAQVAAGQVELSCEV-TGLHRPQITWM--HNTQELGLEEQ 283

QY  357 VRIAVLSDGTLNFTNTVVDGTGMYTCMVNSVG 389
   : ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db  284 TQAEILPFGSLLHRSADTSDMGTYQCIARNEMG 316
```

Search completed: August 26, 2002, 15:47:05
Job time: 147 sec

Result No.	Query	Score	Match	Length	DB	ID	Description
1	430.5	17.1	713	1	GAC1_HUMAN	O75325	homo sapien
2	394	15.6	361	1	CHAD_BOVIN	Q27972	bos taurus
3	356	14.5	567	1	GPV_MOUSE	O08742	mus musculus
4	356	14.1	567	1	GPV_RAT	O08770	rattus norv
5	350	13.9	1480	1	SLIT_DROME	P24014	drosophila
6	329.5	13.1	560	1	GPV_HUMAN	P40197	homo sapien
7	329.5	13.1	605	1	ALS_PAPHA	O02833	papio hamad
8	314.5	12.5	605	1	ALS_HUMAN	P35858	homo sapien
9	309	12.2	603	1	ALS_RAT	P35859	rattus norv
10	302.5	12.0	603	1	ALS_MOUSE	P70389	mus musculus
11	302	12.0	359	1	PGS2_HUMAN	P07585	homo sapien
12	295	11.7	360	1	PGS2_PIG	Q9xsd9	sus scrofa
13	294	11.6	360	1	PGS2_HORSE	O46542	equus cabal
14	293	11.6	360	1	PGS2_BOVIN	P21793	bos taurus
15	291	11.5	357	1	PGS2_CHICK	P28675	gallus gall
16	290	11.5	360	1	PGS2_CANFA	Q29393	canis famil
17	289	11.5	360	1	PGS2_RABIT	Q28888	oryctolagus
18	286.5	11.4	536	1	CBP8_HUMAN	P22792	homo sapien
19	286	11.3	354	1	PGS2_MOUSE	P28654	mus musculus
20	283	11.2	354	1	PGS2_RAT	Q01129	rattus norv
21	282	11.2	626	1	GPBA_HUMAN	P07359	homo sapien
22	276	10.9	353	1	KERA_COTJA	Q9de66	coturnix co
23	274	10.9	353	1	KERA_CHICK	O42235	gallus gall
24	273.5	10.8	331	1	PLIB_AGKBL	O93233	agklistrodon
25	271.5	10.8	368	1	PGS1_HUMAN	P21810	homo sapien
26	271.5	10.8	369	1	PGS1_CANFA	O02678	canis famil
27	271.5	10.8	372	1	PGS1_HORSE	O46403	equus cabal
28	270.5	10.7	369	1	PGS1_MOUSE	P28653	mus musculus
29	270.5	10.7	369	1	PGS1_RAT	P47853	rattus norv
30	269.5	10.7	369	1	PGS1_SHEEP	O46390	ovis aries
31	268.5	10.6	369	1	PGS1_BOVIN	P21809	bos taurus
32	268	10.6	423	1	OMD_RAT	Q921s7	rattus norv
33	259.5	10.3	966	1	Y918_HUMAN	O94991	homo sapien

DR	InterPro: IPR001611; LRR.
DR	InterPro: IPR000483; LRR_Cterm.
DR	InterPro: IPR000372; LRR_Nterm.
DR	InterPro: IPR003592; LRR_Out.
DR	InterPro: IPR003591; LRR_typ.
DR	Pfam: PF00560; LRR; 13.
DR	Pfam: PF01463; LRRCT; 1.
DR	PRINTS: PR00019; LEURICHRPT.
DR	SMART: SM00370; LRR; 1.
DR	SMART: SM00082; LRRCT; 1.
DR	SMART: SM00013; LRRNT; 1.
DR	SMART: SM00369; LRR_TYP; 10.
KW	Platelet; Transmembrane; Glycoprotein; Blood coagulation;
KW	Repeat; Leucine-rich repeat; Cell adhesion; signal.
FT	SIGNAL 1 16 POTENTIAL.
FT	CHAIN 17 567 PLATELET GLYCOPROTEIN V.
FT	DOMAIN 17 522 EXTRACELLULAR (POTENTIAL).
FT	TRANSEM 523 543 POTENTIAL.
FT	DOMAIN 544 567 CYTOPLASMIC (POTENTIAL).
FT	REPEAT 73 96 LRR 1.
FT	REPEAT 97 120 LRR 2.
FT	REPEAT 122 144 LRR 3.
FT	REPEAT 145 168 LRR 4.
FT	REPEAT 170 192 LRR 5.
FT	REPEAT 194 216 LRR 6.
FT	REPEAT 217 240 LRR 7.
FT	REPEAT 241 264 LRR 8.
FT	REPEAT 266 288 LRR 9.
FT	REPEAT 289 312 LRR 10.
FT	REPEAT 314 337 LRR 11.
FT	REPEAT 338 361 LRR 12.
FT	REPEAT 362 385 LRR 13.
FT	REPEAT 386 409 LRR 14.
FT	CARBOHYD 51 51 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 67 67 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 181 181 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 243 243 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 298 298 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 312 312 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 385 385 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ	SEQUENCE 567 AA; 63467 MW; C48643AA73967A7D CRC64;

Query Match 14.5%; Score 367; DB 1; Length 567;
 Best Local Similarity 31.3%; Pred. No. 7e-19;
 Matches 103; Conservative 53; Mismatches 133; Indels 40; Gaps

QY	20	RKNLRVDPGI---	SYNTRLLNHNHQIQIKVNSPKHLRHLLEIQLSRNHRTIIGAF	76
Db	179	RNNLTLPKGLGAQVKLEKLLYSNQLTSVDSGLLSNLTALTELRLRNHLRSVAPGAF	238	
QY	77	NGLANLTLEFNRLTTPINGAFVYLSKLKELWLRNNPTIESPYAFNRIPSLRRDLGL	136	
Db	239	DRGNLSLTLSGNLLESPLPALFVHSSVRSRLTLENPLEEPLDVLFGEMAGRELINL	298	
QY	137	ELKRLSYISGEAFEGLSNRLRYLNLMCMNREIPNLPLIK-----LDELDSGNH---	188	
Db	299	G-THLSLTPAAAFNRLSGTLGLT-----RNPRLSALPRGVFQGLRELRLVGLHTNALA	352	
QY	189	AIRGSPQGLMHLOKLMVIOQTOVERNADFNLQSLVEINLAHNUTLLPHDLFTPLHH	248	
Db	353	ELRDDALRGHLGRVSLRNRLRALPRTLFRNLSSLESVQLEHNQLETLPGDVFAALPQ	412	
QY	249	LERIHLHNPNWCNDTLWLWSWIKWDPASPNTACCARCNTPPNPKRGYIGELDONFTCY	308	
Db	413	LTQVLGHNPWLCDCGLWFREWLRLH-----HPDILGR--DEPPQ-----CR	452	
QY	309	APVIVEPPADLVNTEGMAAEKCRASL	337	
Db	453	GP---EPRASLSPFWELQGDPCDPDRSL	478	

RESULT 4

GPV_RAT ID GPV_RAT STANDARD; PRT; 567 AA.
AC O08770;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Platelet glycoprotein V precursor (GPV) (CD42D).
GN GP5.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; Tissue=Liver;
RX MEDLINE=9275136; PubMed=9129030;
RA Ravanat C., Morales M., Azorsa D.O., Moog S., Schuhler S.,
RA Grunert P., Loew D., van Dorselaer A., Cazenave J.-P., Lanza F.;
RT "Gene cloning of rat and mouse platelet glycoprotein V:
RT identification of megakaryocyte-specific promoters and demonstration
RT of functional thrombin cleavage.";
RL Blood 89:3253-3262(1997).
CC -1- FUNCTION: THE GPIB-V-IX COMPLEX FUNCTIONS AS THE VON WILLEBRAND
CC FACTOR RECEPTOR AND MEDIATES VON WILLEBRAND FACTOR-DEPENDENT
CC PLATELET ADHESION TO BLOOD VESSELS. THE ADHESION OF PLATELETS TO
CC INJURED VASCULAR SURFACES IN THE ARTERIAL CIRCULATION IS A
CC CRITICAL INITIATING EVENT IN HEMOSTASIS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: CONTAINS 14 LEUCINE-RICH REPEATS (LRR).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z69594; CAA93440.1; -
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000483; LRR_Cterm.
DR InterPro: IPR000372; LRR_Nterm.
DR InterPro: IPR003592; LRR_out.
DR InterPro: IPR003591; LRR_typ.
DR Pfam: PF00560; LRR; 14.
DR Pfam: PF01463; LRRCT; 1.
DR PRINTS: PR00019; LEURICHRPT.
DR SMART: SM00370; LRR; 2.
DR SMART: SM00062; LRRCT; 1.
DR SMART: SM00013; LRRNT; 1.
DR SMART: SM00369; LRR_TYP; 10.
KW Platelet; Transmembrane; Glycoprotein; Blood coagulation;
KW Repeat; Leucine-rich repeat; Cell adhesion; Signal.
FT SIGNAL 1 16
FT CHAIN 17 567
FT DOMAIN 17 522
FT TRANSMEM 523 543
FT POTENTIAL.
FT PLATELET GLYCOPROTEIN V.
FT EXTRACELLULAR (POTENTIAL).
FT POTENTIAL.
FT CYTOPLASMIC (POTENTIAL).
FT LRR 1.
FT REPEAT 97 120
FT LRR 2.
FT REPEAT 122 144
FT LRR 3.
FT REPEAT 145 168
FT LRR 4.
FT REPEAT 169 192
FT LRR 5.
FT REPEAT 194 216
FT LRR 6.
FT REPEAT 217 240
FT LRR 7.
FT REPEAT 241 264
FT LRR 8.
FT REPEAT 266 288
FT LRR 9.
FT REPEAT 289 312
FT LRR 10.
FT REPEAT 314 337
FT LRR 11.
FT REPEAT 338 361
FT LRR 12.
FT REPEAT 362 385
FT LRR 13.
FT REPEAT 387 409
FT LRR 14.
FT CARBOHYD 51
FT N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 181 181 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 243 243 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 298 298 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 312 312 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 385 385 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 498 498 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 567 AA; 63344 MW; CA10708E0D03707F CRC64;
Query Match 14.1%; Score 356; DB 1; Length 567;
Best Local Similarity 29.9%; Pred. No. 4.3e-18;
Matches 110; Conservative 53; Mismatches 163; Indels 42; Gaps 9;
QY 6 VCSCSNQFS-----KVICV-RKNLREVDPGI---STNTRLNLNHNQIQI 48
DB 151 LCLNQNLSPANLFSGLKLVLDLSRNLTHLPQGLGCAQIKLEKLLYSNRLSLD 210
QY 49 VNSPKHLRHLLEILOLSRNHRTIEIGAFNGLANLNTLELFDNRLTTIPNGAFVYLSKLE 108
DB 211 SGLLANLGAUTELRLNHLRSIAPGAFDSLGNLSTLTLNGLLESPLPALFLHVSMLTR 270
QY 109 LWLNPNIESIPSAFNRIPLRLDLGELKRLSYISEGAFEGSLNRYLNLANCNLREI 168
DB 271 LTLFENLEELPEVLFGEMAGLRELWNG-THLRTLPAFAAFNLSGLQLGLTRNPL--L 327
QY 169 PNLTP-----LIKLELDLSGNHLSAIRPGSFGQMLHLOKLMWIOSQIQI 223
DB 328 SALPFGMFGHGLTELRLVAVHTNALBELPEDALRGRLRQVSLRHNRLRALPRTFLNLS 387
QY 224 SLVEINLAHNNLTLPLHDLFTPLHLHLRIHLHNPNCNCDILWSMWI-----KDM 275
DB 388 SLVTQLEHNQLKTPGDVFAALPQLRVLLGHNPNWLCDCGLWPFLOWLRHLELLGRDE 447
QY 276 APSWTACARCCTPNKLGRIIGELDQNYFTCYAPVIVEPPADLNVTEGMAAEKLCRAST 335
DB 448 PP-----QCNGPESRASLTFWELQDQMC--PSSRGLPPDPPTNALAKAPDTPRPN 498
QY 336 SLTSSVSWI 343
DB 499 SSQSWAV 506
RESULT 5
SLIT_DROME ID SLIT_DROME STANDARD; PRT; 1480 AA.
AC P24014;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Slit protein precursor.
GN SLI.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91099665; PubMed=2176636;
RA Rothberg J.M., Jacobs J.R., Goodman C.S., Artavanis-Tsakonas S.;
RT "Slit: an extracellular protein necessary for development of midline
RT glia and commissural axon pathways contains both EGF and LRR
RT domains.";
RL Genes Dev. 4:2169-2187(1990).
CC -1- FUNCTION: NECESSARY FOR DEVELOPMENT OF MIDLINE GLIA AND
CC COMMISSURAL AXON PATHWAYS. SLIT MAY INTERACT WITH EXTRACELLULAR
CC MATRIX MOLECULES.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: EXCRETED BY THE MIDLINE GLIA CELLS AND
CC EVENTUALLY DISTRIBUTED ALONG THE AXONS.
CC -1- SIMILARITY: CONTAINS 7 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 24 LEUCINE-RICH REPEATS (LRR).

Matches 83; Conservative 43; Mismatches 104; Indels 49; Gaps 3;

QY 22 NLRVDPG-----ISTNRLN-----LHENOIQIKVNSFKH 54
DB 181 NLTHPLKGLGAQAKLERLLHNSRLVSLDGLNSLGAELTELQPHRNHIRSTAPGAFDR 240
QY 55 LRHLEILQLSRNHIRTIEIGAFNGLANLNTLEFDNRLTTPNGAFYLSKLKELWLRNN 114
DB 241 LPNLSLTLNRHLAFPSALFHLSHNLTLTLFENPLAELPGVLFQGMGLQELWLRNT 300
QY 115 PIESIPYAFNRIPSLRRDLGELKRLSYISEGAFEGLSNLRVNLAMCNLREIPNUTPL 174
DB 301 QLRTLPAAFNRISRLRYLGVTLSPRSALPQGAFOGLGELQV----- 343
QY 175 IKLELDLSNHLSSAIRPGSFQGLMHLQKLMWISQIOVIERNAFDNLQSLVLEINLAHN 234
DB 344 -----LALHNSGLTALPDGLLRLGKLRQVSLRNRRLRALPRALFRNLSSLESVOLDHNO 398
QY 235 LTLPLDPLFTPLHHLRIHJHHPNWCNCDILWLSMWIK 273
DB 399 LETLPGDVFGALPRLTEVLLGHNSWRCDGLGPFPLGLWR 437

RESULT 7

ALS_PAPHA STANDARD; PRT; 605 AA.

AC 002833;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Insulin-like growth factor binding protein complex acid labile chain
DE precursor (ALS).
GN IGFALS OR ALS.
OS Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Papio.
OX NCBI_TaxID=9557;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=97040714; PubMed=8886027;
RA Delhanty P., Baxter R.C.;
RT "The cloning and expression of the baboon acid-labile subunit of the
RT insulin-like growth factor binding protein complex.";
RL Biochem. Biophys. Res. Commun. 227:897-902(1996).
CC -I- FUNCTION: INVOLVED IN PROTEIN-PROTEIN INTERACTIONS THAT RESULT
CC IN PROTEIN COMPLEXES, RECEPTOR-LIGAND BINDING OR CELL ADHESION.
CC -I- SUBUNIT: FORMS A TERNARY COMPLEX OF ABOUT 140 TO 150 KDA WITH
CC IGF-I OR IGF-II AND IGFBP-3 (BY SIMILARITY).
CC -I- SUBCELLULAR LOCATION: Extracellular.
CC -I- SIMILARITY: CONTAINS 20 LEUCINE-RICH REPEATS (LRR).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; S83462; -; NOT_ANNOTATED_CDS.
DR HSSP; P23945; 1XUN.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003592; LRR_Out.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00560; LRR; 19.
DR Pfam; PF01463; LRRCT; 1.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00370; LRR; 2.

DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_TYP; 11.
KW Glycoprotein; Leucine-rich repeat; Repeat; Signal.
FT SIGNAL 1 27
FT CHAIN 28 605
FT REPEAT 53 73
FT REPEAT 74 96
FT REPEAT 98 120
FT REPEAT 121 144
FT REPEAT 146 168
FT REPEAT 169 192
FT REPEAT 194 216
FT REPEAT 217 240
FT REPEAT 242 264
FT REPEAT 265 288
FT REPEAT 289 312
FT REPEAT 313 336
FT REPEAT 338 360
FT REPEAT 361 384
FT REPEAT 386 408
FT REPEAT 409 432
FT REPEAT 433 456
FT REPEAT 458 480
FT REPEAT 482 504
FT REPEAT 505 530
FT CARBOHYD 64 64
FT CARBOHYD 85 85
FT CARBOHYD 96 96
FT CARBOHYD 368 368
FT CARBOHYD 515 515
FT CARBOHYD 580 580
FT SEQUENCE 605 AA; 66110 MW; 9D71994625F23652 CRC64;
Query Match 13.1%; Score 329.5; DB 1; Length 605;
Best Local Similarity 21.0%; Pred. No. 3.5e-16;
Matches 119; Conservative 61; Mismatches 118; Indels 269; Gaps 12;
QY 3 CPVSVCSS-----NQFSKVICVRKNLREVPDGISTNRL----- 36
DB 41 CPATCACSYYDEVNELS-VFCSSRNLRTPDGPGGTQALWLDNSNNLSSIPPAFRNLSS 99
QY 37 -----LNLHENOIQIKVNSFKHLRLETLQLSRNHR 69
DB 100 LAFNLQGGQGLSEPOALLGLENLCHLHRLNQLRSVAVGTFFATALLGLSLNRLS 159
QY 70 TIEIGAFNGLANLNTLEFDNRLTTPNGAF-----VY-----LSK 105
DB 160 RLEDGLFEGGLNLDNLGNWSLAVLPDPAFRGLGLRELVLGNRLAYLQPALFSGLA 219
QY 106 LKELWLRNNPESIPSYAFNRIPSLRL----- 133
DB 220 LRELDLSRNALRAIKANVFAQLPRQLKLYLDRLNIAAVAPGAFGLKALRWLDLSHRVA 279
QY 134 -----DLGE 137
DB 280 GLEDTFGLGLRLVRLSHNAIASLRPRTFEDLHLEELQGLHNRIRQLAERSEFGLGQ 339
QY 138 LK-----RLSYISEGAFEGLSNLRVNLAMCNLREIPN-LTPLIKLELDLSGNHLS 188
DB 340 LEVLTLDHNLQLEVKVGAFLGLTNVAVNNLSGNCLRNLPQEVFGLGKLSLHLEGSCIG 399
QY 189 AIRPGSFQGLMHLQKLMWISQIOVIERNAFDNLQSLVLEINLAHNLTLLPHDLFTPL-- 246
DB 400 RIRPHTFAGLSGLRRLFLKONGLVGIEEQSLWGLAELELDLTSNQLTHPLHQFGLGK 459
QY 247 -----HH----- 248
DB 460 LEYLLLSHNRLAELPADALGPQRAFWDVSHNRLEALPGSLASLGRRLYLNLRNLSLR 519
QY 249 -----LERIHLHNPNCNCDILWLSMWIKMAPSNTACCAR----- 285

```

||||:| |||:|:| | ||:| |
Db 520 TFTPQPPGLERLWLEGNPDWDCSPLKAL-----RDFALQNPSAVPRFVQAICGDDCQPPV 575
QY 286 -----CNTPPNLKGRVIGELDQNYF 305
Db 576 YTYNNITCASPPEVAGLDRLDGEAHF 602

RESULT 8
ALS_HUMAN
ID ALS_HUMAN STANDARD; PRT; 605 AA.
AC P35858;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Insulin-like growth factor binding protein complex acid labile chain
DE precursor (ALS).
GN IGFALS OR ALS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=92357025; PubMed=1379671;
RA Leong S.R., Baxter R.C., Camerato T., Dai J., Wood W.I.;
RT "Structure and functional expression of the acid-labile subunit of
RT the insulin-like growth factor-binding protein complex.";
RL Mol. Endocrinol. 6:870-876(1992).
(2)
RP SEQUENCE FROM N.A.
RA Frankland J.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 28-35.
RX MEDLINE=89308584; PubMed=2473065;
RA Baxter R.C., Martin J.L., Beniac V.A.;
RT "High molecular weight insulin-like growth factor binding protein
RT complex. Purification and properties of the acid-labile subunit from
RT human serum.";
RL J. Biol. Chem. 264:11843-11848(1989).
CC -1- FUNCTION: INVOLVED IN PROTEIN-PROTEIN INTERACTIONS THAT RESULT
CC IN PROTEIN COMPLEXES, RECEPTOR-LIGAND BINDING OR CELL ADHESION.
CC -1- SUBUNIT: FORMS A TERNARY COMPLEX OF ABOUT 140 TO 150 kDa WITH
CC IGF-I OR IGF-II AND IGFBP-3.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: PLASMA.
CC -1- SIMILARITY: CONTAINS 20 LEUCINE-RICH REPEATS (LRR).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M85826; AAA36047.1; -.
DR EMBL; AL031724; CAC36078.1; -.
DR PIR; A41915; A41915.
DR HSP; P23945; 1XUN.
DR MIM; 601489; -.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003592; LRR_Out.
DR InterPro; IPR003591; LRR_Typ.
DR Pfam; PF00560; LRR; 19.
DR Pfam; PF01463; LRRCT; 1.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PR00019; LEURICRPT.
DR SMART; SM00370; LRR; 2.

```

```

DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_Typ; 11.
KW Glycoprotein; Leucine-rich repeat; Repeat; Signal.
FT FT SIGNAL 1 27
FT CHAIN 28 605
FT FT INSULIN-LIKE GROWTH FACTOR BINDING
FT FT PROTEIN COMPLEX ACID LABILE CHAIN.
FT REPEAT 53 73
FT REPEAT 74 96
FT REPEAT 98 120
FT REPEAT 121 144
FT REPEAT 145 168
FT REPEAT 170 192
FT REPEAT 193 216
FT REPEAT 217 240
FT REPEAT 242 264
FT REPEAT 265 288
FT REPEAT 289 312
FT REPEAT 313 336
FT REPEAT 337 360
FT REPEAT 361 384
FT REPEAT 386 408
FT REPEAT 409 432
FT REPEAT 433 456
FT REPEAT 458 480
FT REPEAT 482 504
FT REPEAT 505 530
FT CARBOHYD 64 64
FT CARBOHYD 85 85
FT CARBOHYD 96 96
FT CARBOHYD 368 368
FT CARBOHYD 515 515
FT CARBOHYD 580 580
FT SEQUENCE 605 AA; 66034 MW; F6562A23CB918F6 CRC64;
QY 3 CPVSCSCS-----NQFSKVICVRKLNREVPDGISTNTTLLNLHENQIOIKVNFKHLRH 57
Db 41 CPAACVCSYDDADELS-VFCSSRNLRPLPDGVPGGTQALMDGNLSSVPPAFAFNLS 99
QY 58 LEI-----LQSRNHRTIEIGAFNGLANLNTLEFDNRLT 93
Db 100 LGFLNLOGGQLGLEPOALLGLENLCHLERNQLRSLALGTFAHTPALASGLSNRLS 159
QY 94 TIPNGAFVYLSKELMLRNPNIESIPSYAFNRIPIRRLDLGELKRLSYISEGAFGLS 153
Db 160 RLEDGLFEGLSLWDLNLGMNSLAVLPDAAPRGLSLRELVLG-NRLAYLQPALFSGLA 218
QY 154 NLRYLNLAMCNLRIPN--LTPLIKDELDSLGNHLSAIRPGSQGLMHQLKLMWISQI 211
Db 219 ELREDLSRNALRAIKANVFQLPRLQKLYLDRNLIAAVAPGAFGLKALRMLDLSHRV 278
QY 212 QVIERNAFDNLQSLVEINLAHNLTLLPHDLFTPLHHLERHLLHN 257
Db 279 AGLLDFTFPGLLGLRLVRLSLSHNAIASLRPTFKDLHFLEELQLGHN 324

Query Match 12.5%; Score 314.5; DB 1; Length 605;
Best Local Similarity 31.1%; Pred. No. 4.1e-15;
Matches 89; Conservative 47; Mismatches 117; Indels 33; Gaps 5;
QY 3 CPVSCSCS-----NQFSKVICVRKLNREVPDGISTNTTLLNLHENQIOIKVNFKHLRH 57
Db 41 CPAACVCSYDDADELS-VFCSSRNLRPLPDGVPGGTQALMDGNLSSVPPAFAFNLS 99
QY 58 LEI-----LQSRNHRTIEIGAFNGLANLNTLEFDNRLT 93
Db 100 LGFLNLOGGQLGLEPOALLGLENLCHLERNQLRSLALGTFAHTPALASGLSNRLS 159
QY 94 TIPNGAFVYLSKELMLRNPNIESIPSYAFNRIPIRRLDLGELKRLSYISEGAFGLS 153
Db 160 RLEDGLFEGLSLWDLNLGMNSLAVLPDAAPRGLSLRELVLG-NRLAYLQPALFSGLA 218
QY 154 NLRYLNLAMCNLRIPN--LTPLIKDELDSLGNHLSAIRPGSQGLMHQLKLMWISQI 211
Db 219 ELREDLSRNALRAIKANVFQLPRLQKLYLDRNLIAAVAPGAFGLKALRMLDLSHRV 278
QY 212 QVIERNAFDNLQSLVEINLAHNLTLLPHDLFTPLHHLERHLLHN 257
Db 279 AGLLDFTFPGLLGLRLVRLSLSHNAIASLRPTFKDLHFLEELQLGHN 324

RESULT 9
ALS_RAT
ID ALS_RAT STANDARD; PRT; 603 AA.
AC P35859;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Insulin-like growth factor binding protein complex acid labile chain
DE precursor (ALS).
GN IGFALS OR ALS.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=93038676; PubMed=1384485;
RA Dai J., Baxter R.C.;
RT "Molecular cloning of the acid-labile subunit of the rat insulin-like
growth factor binding protein complex.";
RL Biochem. Biophys. Res. Commun. 188:304-309(1992).
RN [2]
RP SEQUENCE OF 24-44, AND CHARACTERIZATION.
RC STRAIN=Wistar; TISSUE=Serum;
RX MEDLINE=94130835; PubMed=7507839;
RA Baxter R.C., Dai J.;
RT "Purification and characterization of the acid-labile subunit of rat
serum insulin-like growth factor binding protein complex.";
RL Endocrinology 134:848-852(1994).
CC -1- FUNCTION: MAY HAVE AN IMPORTANT ROLE IN REGULATING THE ACCESS OF
CIRCULATING IGFs TO THE TISSUES.
CC -1- SUBUNIT: FORMS A TERNARY COMPLEX OF ABOUT 140 TO 150 kDa WITH
IGF-I OR IGF-II AND IGFBP-3.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: BRAIN, KIDNEY, LUNG, HEART, SPLEEN, MUSCLE
AND LIVER.
CC -1- SIMILARITY: CONTAINS 21 LEUCINE-RICH REPEATS (LRR).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
DR EMBL: S46785; AAB33770.2; ..
DR PIR: JCI282; JCI282.
DR HSP; P23945; 1XUN.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000483; LRR.Cterm.
DR InterPro: IPR000372; LRR.Nterm.
DR InterPro: IPR003592; LRR.Out.
DR InterPro: IPR003591; LRR_Typ.
DR Pfam: PF00560; LRR; 19.
DR Pfam: PF01463; LRRCT; 1.
DR Pfam: PF01462; LRRNT; 1.
DR PRINTS: PR00019; LEURICHRPT.
DR SMART: SM00370; LRR; 5.
DR SMART: SM00082; LRRCT; 1.
DR SMART: SM00013; LRRNT; 1.
DR SMART: SM00369; LRR_Typ; 9.
KW Glycoprotein; Leucine-rich repeat; Repeat; Signal.
FT SIGNAL 1 23
FT CHAIN 24 603 INSULIN-LIKE GROWTH FACTOR BINDING
FT PROTEIN COMPLEX ACID LABILE CHAIN.
FT REPEAT 52 73
FT REPEAT 74 96
FT REPEAT 98 120
FT REPEAT 121 144
FT REPEAT 146 168
FT REPEAT 169 192
FT REPEAT 194 216
FT REPEAT 217 240
FT REPEAT 242 264
FT REPEAT 266 288
FT REPEAT 289 312
FT REPEAT 313 336
FT REPEAT 337 360
FT REPEAT 361 384
FT REPEAT 386 408
FT REPEAT 409 432
FT REPEAT 433 456
FT REPEAT 458 478
FT REPEAT 478 500

FT REPEAT 479 504 LRR 19.
FT REPEAT 506 529 LRR 20.
FT REPEAT 543 566 LRR 21.
FT CARBOHYD 64 64 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 85 85 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 515 515 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 578 578 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 586 586 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 603 AA; 66811 MW; DCD7637D94A5037C CRC64;
Query Match 12.2%; Score 309; DB 1; Length 603;
Best Local Similarity 30.3%; Pred. No. 1e-14;
Matches 96; Conservative 43; Mismatches 126; Indels 52; Gaps 8;
QY 36 LNLH-----ENQIIVKNSFKHLRHLLEILQLSRNIHRTIEIGAFNGLANLNFLEFDN 90
DB 289 LGLHLVRLAHHNAIASLRPTFKDLHLEELQLGHNRIRQLGERTFEGLGQLEVLTLND 348
QY 91 RLTPNGAFVYLSKELMLRNPIESIPSYAFNRIPSLRRDLGLKRLSYISGAFE 150
DB 349 QITEVRVGAESGLFNVAVMNLSGCLRSPLRVFQGLKHLHL-EHSCIGHVRLHTFA 407
QY 151 GLSNRLYNLAMCNLREI--PNLTPLIKDELDELDSGHLNLSAIRPGSQGLMHLQKLMWQ 208
DB 408 GLSGLRLFLRDSNISIEQSGLSELLELDLTTNRLTHLPQLFGHLEVLILSY 467
QY 209 SQIOVERNAFDNLQSLVEINLAHNNLTLLPHDL-----FTPLHHL 249
DB 468 NQLTTLTSAEVLGPIQAFWLDISHNHLTAEGLFSSIGRVRVYLSLRNNSIQTFSPQGL 527
QY 250 ERTHLHNPNCNDILLSWKIDMAPSN-----TAC-----ARCNT 288
DB 528 ERLWLDANPWCDCPLKAL----RDFALQNPVVPRVQTVCEGDGCPVYNNITCAG 583
QY 289 PPNLKGRYIGELQNYF 305
DB 584 PANVSGDLDRDVSETHF 600
RESULT 10
ALS_MOUSE
ID ALS_MOUSE STANDARD; PRT; 603 AA.
AC P70389;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Insulin-like growth factor binding protein complex acid labile chain
DE precursor (ALS).
GN IGFALS OR ALS OR ALBS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RX MEDLINE=96413591; PubMed=8816745;
RA Boisclair Y.R., Seto D., Hsieh S., Hurst K.R., Ooi G.T.;
RT "Organization and chromosomal localization of the gene encoding the
mouse acid labile subunit of the insulin-like growth factor binding
complex.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:10028-10033(1996).
CC -1- FUNCTION: MAY HAVE AN IMPORTANT ROLE IN REGULATING THE ACCESS OF
CIRCULATING IGFs TO THE TISSUES.
CC -1- SUBUNIT: FORMS A TERNARY COMPLEX OF ABOUT 140 TO 150 kDa WITH
IGF-I OR IGF-II AND IGFBP-3 (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: CONTAINS 21 LEUCINE-RICH REPEATS (LRR).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; U66900; AAB17270.1; -

MGD; MGI:107973; Igfals.

InterPro; IPR001611; LRR.

InterPro; IPR000483; LRR_Cterm.

InterPro; IPR000372; LRR_Nterm.

InterPro; IPR003592; LRR_Out.

InterPro; IPR003591; LRR_Typ.

Pfam; PF00560; LRR; 19.

Pfam; PF01463; LRRCT; 1.

Pfam; PF01462; LRRNT; 1.

PRINTS; PR00019; LEURICHRPT.

SMART; SM00370; LRR; 4.

SMART; SM00082; LRRCT; 1.

SMART; SM00013; LRRNT; 1.

SMART; SM00369; LRR_Typ; 9.

Glycoprotein; leucine-rich repeat; Repeat; Signal.

BY SIMILARITY.

INSULIN-LIKE GROWTH FACTOR BINDING

PROTEIN COMPLEX ACID LABILE CHAIN.

REPEAT 52 73

REPEAT 74 96

REPEAT 98 120

REPEAT 121 144

REPEAT 146 168

REPEAT 169 192

REPEAT 194 216

REPEAT 217 240

REPEAT 242 264

REPEAT 265 288

REPEAT 289 312

REPEAT 313 336

REPEAT 338 360

REPEAT 361 384

REPEAT 386 408

REPEAT 409 432

REPEAT 433 456

REPEAT 458 480

REPEAT 482 504

REPEAT 505 529

REPEAT 543 566

REPEAT 64 64

CARBOHYD 85 85

CARBOHYD 96 96

CARBOHYD 368 368

CARBOHYD 515 515

CARBOHYD 578 578

CARBOHYD 586 586

SEQUENCE 603 AA; 66959 MW; 4A6223ABC7026BCC CRC64;

Query Match 12.0%; Score 302.5; DB 1; Length 603;

Best Local Similarity 31.5%; Pred. No. 2.9e-14;

Matches 93; Conservative 41; Mismatches 110; Indels 51; Gaps 7;

QY 3 CPVCTSCS-----NQFSKVICVRKNLRVDPGISTNTRLLNLHNOIQIKVNSFKHLRH 57

DB 41 CPVCTSCSYDDYDLS-VFCCSRNLQLPDGPVSTRALWLDGNLSSIPSAFQNLSS 99

QY 58 LEILQLSRHRTTEICAFNGLANLNTLEPDNRLTTPNGAFVLSKLKELWRNPPIE 117

DB 100 LDFNLQGSWRLEPQALLGLQLNLYHLHRLNRLSLAAGLFRHTPSLASLSGLNLLG 159

QY 118 SIPSFAFNRIPLSRDLGELKRLSYISEGAFGLSNLRVNLTA-----MCNLR 166

DB 160 RLEGLFRGLSHLDNLG-WNSLVLPDVTVFQGLGNLHELVLGAGNKLTYLPALLCLGL 218

QY 167 EIPNLTPLIKDELDSNLHLSAIRPGSFOGLMHLQKLMMIQSIQVIERNAF----- 219

DB 219 E-----LRELDLSRNALRSYKANVFTHLPRLQKLYLDRLNLTAVAPRAFLGMKALR 269

QY 220 -----DNLOSLVE-----INLAHNNLTLLPHDLETPPLHLHLRIHLHN 257

DB 270 WLDLSHNRVAGLLEDTFPGLLGLHVLRLAHNAITSLRPTFKDLHFLLEQLQGHN 324

RESULT 11

PGS2_HUMAN

ID PGS2_HUMAN STANDARD; PRT; 359 AA.

AC P07585; Q9Y5N8; Q9P020; Q9P021;

DT 01-APR-1988 (Rel. 07, Created)

DT 01-APR-1988 (Rel. 07, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)

DE Bone proteoglycan II precursor (PG-S2) (Decorin) (PG40).

GN DCN.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID:9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=87017013; PubMed=3484330;

RA Krustus T., Ruoslahti E.;

RT "Primary structure of an extracellular matrix proteoglycan core

protein deduced from cloned cDNA.";

RL Proc. Natl. Acad. Sci. U.S.A. 83:7683-7687(1986).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=93162643; PubMed=8432527;

RA Vetter U., Vogel W., Just W., Young M.F., Fisher L.W.;

RT "Human decorin gene: intron-exon junctions and chromosomal

localization.";

RL Genomics 15:161-168(1993).

RN [3]

RP SEQUENCE OF 1-70 FROM N.A.

RX MEDLINE=93162642; PubMed=8432526;

RA Danielson K.G., Fazio A., Cohen I.R., Cannizzaro L., Iozzo R.V.;

RT "The human decorin gene: intron-exon organization, discovery of two

alternatively spliced exons in the 5' untranslated region, and

mapping of the gene to chromosome 12q23.";

RL Genomics 15:146-160(1993).

RN [4]

RP SEQUENCE FROM N.A. (ISOFORMS A; B; C; D AND E).

RA CS-Szabo G., Giant T.T.;

RT "Alternative splicing of human decorin.";

RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

RN [5]

RP SEQUENCE OF 31-50.

RX MEDLINE=90073579; PubMed=2590169;

RA Roughley P.J., White R.J.;

RT "Permatan sulphate proteoglycans of human articular cartilage. The

properties of dermatan sulphate proteoglycans I and II.";

RL Biochem. J. 262:823-827(1989).

RN [6]

RP SEQUENCE OF 31-49.

RX MEDLINE=87250639; PubMed=3597437;

RA Fisher L.W., Hawkins G.R., Tuross N., Termini J.D.;

RT "Purification and partial characterization of small proteoglycans I

and II, bone sialoproteins I and II, and osteonectin from the mineral

compartment of developing human bone.";

RL J. Biol. Chem. 262:9702-9708(1987).

CC -1- FUNCTION: BINDS TO TYPE I AND TYPE II COLLAGEN AND AFFECTS THE

RATE OF FIBRILS FORMATION. ALSO BINDS TO FIBRONECTIN AND TGF-

BETA.

CC -1- PTM: THE GLYCOSAMINOGLYCAN CHAIN ATTACHED TO DECORIN CAN BE EITHER

CHONDROITIN SULFATE OR DERMATAN SULFATE DEPENDING UPON THE

TISSUE OF ORIGIN.

CC -1- ALTERNATIVE PRODUCTS: 5 isoforms; A (shown here), B, C, D and E;

are produced by alternative splicing.

CC

QY 123 AFNRIPSLRLDLG--ELKRLSYISEGAFGLSNRLYNLNCMLNREI-PNLTPLIKLE 179
 DB 170 VFNGLNQIMIVELGTNPLKS-SGIENGAFQGMKKLSYIRIADNTNITIPGLPP--SLTE 226
 QY 180 LDLSGNHLSAIRPGSFOGLMHLOKLMWISQIQIUVIERNAFDNLQSLVEINLAHNNLTLLP 239
 DB 227 LHLGDKNKTIVDAASLGLNLAAGLGLSFSISAVDNGSLANTPHLRHLNDNNKLIKVP 286
 QY 240 HDLFTPLHLERHLHNPNCNCDILWLSWIKDMAPSNACCARCNTPPNKGRIYIG- 298
 DB 287 GGL-ADHKYIQVYVLLHN--NISAI-----GSDNFCPPGYNTK---KASYGV 328
 QY 299 -----ELDQNYFTC 307
 DB 329 SLFSNPVOYWEIQPSTERC 347

RESULT 14
 PGS2_BOVIN STANDARD; PRT; 360 AA.
 AC P21793;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Bone proteoglycan II precursor (PG-S2) (Decorin).
 GN DCN.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88133946; PubMed=3435485;
 RA Day A.A., McQuillan C.I., Termine J.D., Young M.R.;
 RT "Molecular cloning and sequence analysis of the cDNA for small
 RT proteoglycan II of bovine bone."
 RL Biochem. J. 248:801-805(1987).
 RN [2]
 RP SEQUENCE OF 31-54.
 RX MEDLINE=89123388; PubMed=2914936;
 RA Choi H.U., Johnson T.L., Pal S., Tang L.H., Rosenberg L., Neame P.J.;
 RT "Characterization of the dermatan sulfate proteoglycans, DS-PGI and
 RT DS-PGII, from bovine articular cartilage and skin isolated by octyl-
 RT sepharose chromatography."
 RL J. Biol. Chem. 264:2876-2884(1989).
 CC -1- FUNCTION: BINDS TO TYPE I AND TYPE II COLLAGEN AND AFFECTS THE
 CC RATE OF FIBRILS FORMATION. ALSO BINDS TO FIBRONECTIN AND TGF-
 CC BETA.
 CC -1- PTM: THE GLYCOSAMINOGLYCAN CHAIN ATTACHED TO DECORIN CAN BE EITHER
 CC CHONDROITIN SULFATE OR DERMATAN SULFATE DEPENDING UPON THE
 CC TISSUE OF ORIGIN.
 CC -1- SIMILARITY: BELONGS TO THE SMALL INTERSTITIAL PROTEOGLYCANS
 CC FAMILY.
 CC -1- SIMILARITY: CONTAINS 10 LEUCINE-RICH REPEATS (LRR).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Y00712; CAA68702.1; --
 CC PIR: S06280; S06280.
 CC PIR: B31430; B31430.
 CC InterPro: IPR001611; LRR.
 CC InterPro: IPR000372; LRR.Nterm.
 CC InterPro: IPR003592; LRR.Out.
 CC InterPro: IPR003591; LRR.Typ.
 CC Pfam: PF00560; LRR; 9.

DR Pfam: PF01462; LRRNT; 1.
 DR SMART: SM00370; LRR; 2.
 DR SMART: SM0013; LRRNT; 1.
 DR SMART: SM00369; LRR_Typ; 2.
 KW Glycoprotein; Connective tissue; Extracellular matrix; Proteoglycan;
 KW Repeat; Leucine-rich repeat; Signal.
 FT SIGNAL 1 16
 FT PROPEP 17 30
 FT CHAIN 31 360 BONE PROTEOGLYCAN II.
 FT REPEAT 78 99 LRR 1.
 FT REPEAT 100 123 LRR 2.
 FT REPEAT 124 146 LRR 3.
 FT REPEAT 147 168 LRR 4.
 FT REPEAT 169 194 LRR 5.
 FT REPEAT 195 218 LRR 6.
 FT REPEAT 219 239 LRR 7.
 FT REPEAT 240 263 LRR 8.
 FT REPEAT 264 286 LRR 9.
 FT REPEAT 287 309 LRR 10.
 FT CARBOHYD 34 34 O-LINKED (GLYCOSAMINOGLYCAN) (BY
 FT SIMILARITY).
 FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DISULFID 55 68 BY SIMILARITY.
 FT DISULFID 314 347 BY SIMILARITY.
 SQ SEQUENCE 360 AA; 39837 MW; 71E84DA2D87552C0 CRC64;

Query Match 11.6%; Score 293; DB 1; Length 360;
 Best Local Similarity 29.9%; Pred. No. 7.2e-14;
 Matches 95; Conservative 46; Mismatches 139; Indels 38; Gaps 10;

QY 3 CPVSCSCSNQFSKVICVRKLNREVPDGIQSTNRLNLHENOIQIIVKNSFKHLRHLILQ 62
 DB 55 CPFRQC--HLRVQCSDLGLKVPKDLPPDTALLDQNNKIYKIDGDFKNLKNLTLI 112
 QY 63 LSRNHRTIEIGAFNGLIANTLLELFDNRLTTPNGAFVYLSKLKELWLRNNTIESIPSY 122
 DB 113 LINNKISKISPGAFAPLVKLERLYLSKNQKELPE---KMPKTLQELRVHENEITKVRKS 169
 QY 123 AFNRIPSLRLDLG--ELKRLSYISEGAFGLSNRLYNLNCMLNREI-PNLTPLIKLE 180
 DB 170 VFNGLNQIMIVELGTNPLKS-SGIENGAFQGMKKLSYIRIADNTNITIPGLP-PSLTEL 227
 QY 181 LDLSGNHLSAIRPGSFOGLMHLOKLMWISQIQIUVIERNAFDNLQSLVEINLAHNNLTLLPH 240
 DB 228 HLDGKTKIVDAASLGLNLAAGLGLSFSISAVDNGSLANTPHLRHLNDNNKLIKVP 287
 QY 241 DLFTPLHLERHLHNPNCNCDILWLSWIKDMAPSNACCARCNTPPNKGRIYIG-- 298
 DB 288 GV-ADHKYIQVYVLLHN--NISAI-----GSDNFCPPGYNTK---KASYGSVS 329
 QY 299 -----ELDQNYFTC 307
 DB 330 LFSNPVOYWEIQPSTERC 347

RESULT 15
 PGS2_CHICK STANDARD; PRT; 357 AA.
 ID PGS2_CHICK
 AC P28675;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Bone proteoglycan II precursor (PG-S2) (Decorin).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RC STRAIN-WHITE LEGHORN; TISSUE=Cornea;
RX MEDLINE=92256755; PubMed=1605630;
Li W., Vergnes J.P., Cornuet P.K., Hassel J.R.;
RT "cDNA clone to chick corneal chondroitin/dermatan sulfate
RT proteoglycan reveals identity to decorin.";
RL Arch. Biochem. Biophys. 296:190-197(1992).
CC -|- FUNCTION: BINDS TO TYPE I AND TYPE II COLLAGEN AND AFFECTS THE
CC RATE OF FIBRILS FORMATION. ALSO BINDS TO FIBRONECTIN AND TGF-
CC BETA.
CC -|- PTM: THE GLYCOSAMINOGLYCAN CHAIN ATTACHED TO DECORIN CAN BE EITHER
CC CHONDROITIN SULFATE OR DERMATAN SULFATE DEPENDING UPON THE
CC TISSUE OF ORIGIN (BY SIMILARITY).
CC -|- SIMILARITY: BELONGS TO THE SMALL INTERSTITIAL PROTEOGLYCANS
CC FAMILY.
CC -|- SIMILARITY: CONTAINS 10 LEUCINE-RICH REPEATS (LRR).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X63797; CAA45318.1; -
CC PIR; S22197; S22197.
CC PIR; S24317; S24317.
CC InterPro; IPR001611; LRR.
CC InterPro; IPR000372; LRR_Nterm.
CC InterPro; IPR003592; LRR_Out.
CC InterPro; IPR003591; LRR_typ.
CC Pfam; PF00560; LRR; 10.
CC Pfam; PF01462; LRRNT; 1.
CC SMART; SM00370; LRR; 2.
CC SMART; SM00013; LRRNT; 1.
CC SMART; SM00369; LRR_typ; 2.
CC Glycoprotein; Connective tissue; Extracellular matrix; Proteoglycan;
CC Repeat; Leucine-rich repeat; Signal.
FT SIGNAL 1 16
FT PROPEP 17 30 BONE PROTEOGLYCAN II.
FT CHAIN 31 357 LRR 1.
FT REPEAT 75 96 LRR 2.
FT REPEAT 97 120 LRR 3.
FT REPEAT 121 143 LRR 4.
FT REPEAT 144 165 LRR 5.
FT REPEAT 166 191 LRR 6.
FT REPEAT 192 215 LRR 7.
FT REPEAT 216 236 LRR 8.
FT REPEAT 237 260 LRR 9.
FT REPEAT 261 283 LRR 10.
FT REPEAT 284 306 O-LINKED (GLYCOSAMINOGLYCAN) (BY
FT CARBOHYD 34 SIMILARITY).
FT CARBOHYD 209 209 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 260 260 N-LINKED (GLCNAC...) (POTENTIAL).
FT DISULFID 52 65 BY SIMILARITY.
FT DISULFID 311 344 BY SIMILARITY.
SQ SEQUENCE 357 AA; 39687 MW; 31B104C7C3CD711D CRC64;

Query Match 11.5%; Score 291; DB 1; Length 357;
Best Local Similarity 33.1%; Pred. No. 9.9e-14;
Matches 85; Conservative 37; Mismatches 125; Indels 10; Gaps 6;
QY 3 CPVSVCSCNDFKVCYRKNLREVPDGIISNTRLLNLHENOIQIIVNSFKHLRHLEILQ 62
Db 52 CPFFQCQ--HLRVVQCSDLGLERVKDPDPTLLDLDQNNKITEKEGDFKKNLKHALI 109
QY 63 LSRNHRTIEGAFGLANLNTLELFDNLTTPNGAFVYLSKLKELWLRNRPTEPSY 122
Db 110 LVNNKISKISPAAPLKKLRLYLSKNLKPEN---MPKSLOEIRAHENEISKLRKA 166
QY 123 AFNRIPSRLRLDGL--ELKRLSYISEGAFGLSNLRYLNLMCNLREIPNLTPLIKLDEL 180

Db 167 VFENGLNQVIVLELGTNPLKS-SGIENGAFQGMKRLSVIRIADTNITSPKGLP-PSLTEL 224
QY 181 DLSGNHLISAIRPGSFQGLMHLQKLWMTQSQIQVIERNAFONLQSLVEINLAHNNLTLLPH 240
Db 225 HLDGNKISKIDAEGLSGLTNLAKLGLSFNSISVSENGSLNNVPHLREHLNNNELVRVPS 284
QY 241 DLFTPLHLHLRIHLHN 257
Db 285 GL-GEHKYIQVVYLHN 300
Search completed: August 26, 2002, 15:50:43
Job time: 240 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 15:46:18 ; Search time 31.39 Seconds
(without alignments)
2634.329 Million cell updates/sec

Title: US-09-905-056-292_COPY_45_522

Perfect score: 2524

Sequence: 1 OTCPGVCSNQNFSKVICVR.....KTFITPVTINDSINGIPIDEV 478

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2524	100.0	640	4	Q9HCJ2
2	1581.5	62.7	653	4	Q9HBW1
3	1556.5	61.7	649	4	Q96A85
4	1151.5	45.6	422	4	Q9NT99
5	872.5	34.6	441	11	Q99PH1
6	565.5	22.4	606	4	Q9BZ20
7	565.5	22.4	614	4	Q96FE5
8	560.5	22.2	614	6	Q9N008
9	559.5	22.2	614	11	Q901T0
10	453	17.9	719	4	Q96NI6
11	443.5	17.6	1021	5	Q9V430
12	421	16.7	705	4	Q43377
13	421	16.7	708	4	Q9H3W5
14	413.5	16.4	1093	4	Q96JA1
15	413.5	16.4	1094	4	Q9BYB8
16	410.5	16.3	730	4	Q9P231

17	410	16.2	707	11	P97860
18	408	16.2	789	6	Q9BE71
19	406	16.1	832	4	Q9ULH4
20	403	16.0	788	11	Q9CYK3
21	402.5	15.9	1091	11	P70193
22	401.5	15.9	359	4	Q96RJ5
23	401.5	15.9	492	11	Q99KT6
24	398.5	15.8	716	11	O61809
25	398	15.8	718	13	O73675
26	397.5	15.7	359	4	O15335
27	396	15.7	707	11	Q9ESY6
28	393.5	15.6	1531	11	Q9WVB5
29	389.5	15.4	1515	13	Q9DE37
30	389.5	15.4	1531	11	O88279
31	389	15.4	1521	11	O9RLB9
32	388.5	15.4	358	11	O55226
33	388.5	15.4	674	4	O9NZ01
34	388	15.4	1521	4	O95710
35	388	15.4	1529	4	O94813
36	387.5	15.4	1530	11	O9WUG5
37	386	15.3	1525	4	O9Y507
38	384.5	15.2	1496	4	O92626
39	382.5	15.2	358	11	O70210
40	380	15.1	628	4	O9BTN0
41	378	15.0	796	11	Q9WVC1
42	377.5	15.0	1173	5	O9V7J8
43	377	14.9	1534	4	O75093
44	376	14.9	1065	4	O94898
45	375	14.9	660	4	O43155

ALIGNMENTS

RESULT 1

Q9HCJ2	ID	Q9HCJ2	PRELIMINARY;	PRT;	640 AA.
AC	Q9HCJ2;				
DT	01-MAR-2001 (TrEMBLrel. 16, Created)				
DT	01-MAR-2001 (TrEMBLrel. 16, Last sequence update)				
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)				
DE	KIAA1580 PROTEIN (FRAGMENT).				
GN	KIAA1580.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=20450683; PubMed=10997877;				
RA	Nagase T., Kikuno R., Nakayama M., Hirotsawa M., Ohara O.;				
RT	"Prediction of the coding sequences of unidentified human genes.				
RT	XVIII. The complete sequences of 100 new cDNA clones from brain which				
RT	code for large proteins in vitro."				
RL	DNA Res. 7:273-281(2000).				
DR	EMBL; AB046800; BAB13406.1; -				
DR	HSSP; P22888; ILUT.				
DR	InterPro; IPR003599; Ig.				
DR	InterPro; IPR003598; Ig_c2.				
DR	InterPro; IPR003600; Ig_like.				
DR	InterPro; IPR003006; Ig_MHC.				
DR	InterPro; IPR001611; LRR.				
DR	InterPro; IPR000483; LRR_Cterm.				
DR	InterPro; IPR000372; LRR_Nterm.				
DR	InterPro; IPR003592; LRR_out.				
DR	InterPro; IPR003591; LRR_typ.				
DR	Pfam; PF00047; Ig; 1.				
DR	Pfam; PF00560; LRR; 9.				
DR	Pfam; PF01463; LRRCT; 1.				
DR	Pfam; PF01462; LRRNT; 1.				
DR	PRINTS; PR00019; LEURICHRPT.				
DR	SMART; SM00409; IG; 1.				
DR	SMART; SM00408; IGC2; 1.				

p97860	mus musculus
q9be71	macaca fasc
q9ulh4	homo sapien
q9cyk3	mus musculus
p70193	mus musculus
q96rj5	homo sapien
q99kt6	mus musculus
o61809	mus musculus
o73675	xenopus lae
o15335	homo sapien
q9esy6	rattus norv
q9wvb5	mus musculus
q9de37	brachydanio
o88279	rattus norv
q9rlb9	mus musculus
o55226	mus musculus
q9nzul	homo sapien
o95710	homo sapien
o94813	homo sapien
q9wug5	rattus norv
q9y507	homo sapien
q92626	homo sapien
o70210	rattus norv
q9btn0	homo sapien
q9wvc1	rattus norv
q9v7j8	drosophila
o75093	homo sapien
o94898	homo sapien
o43155	homo sapien

```
DR SMART; SM00410; IG_like; 1.
DR SMART; SM00370; LRR; 6.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_TTP; 8.
KW Immunoglobulin domain.
FT NON_TER 1
SQ SEQUENCE 640 AA; 71949 MW; 6058974872636838 CRC64;

Query Match 100.0%; Score 2524; DB 4; Length 640;
Best Local Similarity 100.0%; Pred. No. 1.8e-181;
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QTCPSVCSNQFSKVCVRKLNREVPDGI STNRLNLHENQIQI IKVNSFKHLRLEI 60
DB 45 QTCPSVCSNQFSKVCVRKLNREVPDGI STNRLNLHENQIQI IKVNSFKHLRLEI 104

QY 61 LQLSRNHRTIEIGAFNGLANLTLELFDNRLTTPNGAFVYLSKLKELWLRNPNIESIP 120
DB 105 LQLSRNHRTIEIGAFNGLANLTLELFDNRLTTPNGAFVYLSKLKELWLRNPNIESIP 164

QY 121 SYAFNRPSLRRLDGLGELKRLSYISEGAFGLSNRLYLNLAMCNLREIPNLTPLIKLDEL 180
DB 165 SYAFNRPSLRRLDGLGELKRLSYISEGAFGLSNRLYLNLAMCNLREIPNLTPLIKLDEL 224

QY 181 DLSGNHLSAIRPGSFQGLMHQKLMQISOQVIERNAFDNLQSLVEINLAHNLTLLPH 240
DB 225 DLSGNHLSAIRPGSFQGLMHQKLMQISOQVIERNAFDNLQSLVEINLAHNLTLLPH 284

QY 241 DLFTPLHLHLRIHLHHPNWCNCDILWLSWKIDMAPSNTACCRCNTPPNPKGRYIGEL* 300
DB 285 DLFTPLHLHLRIHLHHPNWCNCDILWLSWKIDMAPSNTACCRCNTPPNPKGRYIGEL 344

QY 301 DONFTCYAPVIVEPPADLVNTEGMAELKCRASSTLSVSWITPNGTVMTHGAYKVRIA 360
DB 345 DONFTCYAPVIVEPPADLVNTEGMAELKCRASSTLSVSWITPNGTVMTHGAYKVRIA 404

QY 361 VLSGDTLNFTNVTVDGTGMYTCMVNSVGNNTASATLNVTAAATTPFSYFSTVTVETMEP 420
DB 405 VLSGDTLNFTNVTVDGTGMYTCMVNSVGNNTASATLNVTAAATTPFSYFSTVTVETMEP 464

QY 421 SQDEARTDNNVGPVVDWETNVTTSLPQSTRSTEKTFTIPVTDINSIGPIDEV 478
DB 465 SQDEARTDNNVGPVVDWETNVTTSLPQSTRSTEKTFTIPVTDINSIGPIDEV 522

RESULT 2
Q9HBW1 PRELIMINARY; PRT; 653 AA.
AC Q9HBW1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE BRAIN TUMOR ASSOCIATED PROTEIN NAG14.
GN NAG14.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang J., Bin L., Jiang N., Li G.;
RT "Homo sapiens brain-specific gene (BAG), downregulated in brain tumor, mRNA.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF196976; AAC28019.2; -
DR HSSP; P23945; 1XUN.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001611; LRR.
```

```
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003592; LRR_Out.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00560; LRR; 8.
DR Pfam; PF01462; LRRCT; 1.
DR Pfam; PF01463; LRRNT; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00408; IGc2; 1.
DR SMART; SM00410; IG_like; 1.
DR SMART; SM00370; LRR; 3.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_TTP; 7.
KW Immunoglobulin domain.
SQ SEQUENCE 653 AA; 72717 MW; 38159C81F6850E37 CRC64;

Query Match 62.7%; Score 1581.5; DB 4; Length 653;
Best Local Similarity 60.9%; Pred. No. 1.3e-110;
Matches 298; Conservative 77; Mismatches 93; Indels 21; Gaps 8;

QY 1 QTCPSVCSNQFSKVCVRKLNREVPDGI STNRLNLHENQIQI IKVNSFKHLRLEI 60
DB 44 QTCPSVCSNQFSKVCVRKLNREVPDGI STNRLNLHENQIQI IKVNSFKHLRLEI 103

QY 61 LQLSRNHRTIEIGAFNGLANLTLELFDNRLTTPNGAFVYLSKLKELWLRNPNIESIP 120
DB 104 LQLSRNHRTIEIGAFNGLANLTLELFDNRLTTPNGAFVYLSKLKELWLRNPNIESIP 163

QY 121 SYAFNRPSLRRLDGLGELKRLSYISEGAFGLSNRLYLNLAMCNLREIPNLTPLIKLDEL 180
DB 164 SYAFNRPSLRRLDGLGELKRLSYISEGAFGLSNRLYLNLAMCNLREIPNLTPLIKLDEL 223

QY 181 DLSGNHLSAIRPGSFQGLMHQKLMQISOQVIERNAFDNLQSLVEINLAHNLTLLPH 240
DB 224 DLSGNHLSAIRPGSFQGLMHQKLMQISOQVIERNAFDNLQSLVEINLAHNLTLLPH 283

QY 241 DLFTPLHLHLRIHLHHPNWCNCDILWLSWKIDMAPSNTACCRCNTPPNPKGRYIGEL 300
DB 284 DLFTPLHLHLRIHLHHPNWCNCDILWLSWKIDMAPSNTACCRCNTPPNPKGRYIGEL 343

QY 301 DONFTCYAPVIVEPPADLVNTEGMAELKCRASSTLSVSWITPNGTVMTHGAYKVRIA 360
DB 344 DONFTCYAPVIVEPPADLVNTEGMAELKCRASSTLSVSWITPNGTVMTHGAYKVRIA 402

QY 361 VLSGDTLNFTNVTVDGTGMYTCMVNSVGNNTASATLNVTAAATTPFSYFSTVTVETMEP 418
DB 403 VLSGDTLNFTNVTVDGTGMYTCMVNSVGNNTASATLNVTAAATTPFSYFSTVTVETMEP 462

QY 419 EPSQDEARTDNNVGPVVDWETNVTTSLPQSTRSTEKTFTIPVTDINSIGPIDEV 470
DB 463 EPSQDEARTDNNVGPVVDWETNVTTSLPQSTRSTEKTFTIPVTDINSIGPIDEV 513

QY 471 GI-PGIDEV 478
DB 514 KMQTSLDEV 522

RESULT 3
Q96A85 PRELIMINARY; PRT; 649 AA.
ID Q96A85;
AC Q96A85;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 72.3 KDA PROTEIN.
GN BAG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
```



```
DR InterPro; IPR003591; LRR_Typ.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00560; LRR; 2.
DR Pfam; PF01463; LRRCT; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00410; IG_Like; 1.
DR SMART; SM00370; LRR; 2.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00369; LRR_Typ; 3.
KW Immunoglobulin domain.
FT NON_TER 1
SQ SEQUENCE 441 AA; 48844 MW; E2D17AB4607AC475 CRC64;

Query Match 34.68; Score 872.5; DB 11; Length 441;
Best Local Similarity 52.28; Pred. No. 1.4e-57;
Matches 163; Conservative 57; Mismatches 87; Indels 5; Gaps 4;

Qy 170 NLTPLIKDELDSGNHLSAIRPGSFGGLMHLQKLWMIQSQIQVIERNAFNLQSLVEIN 229
Db 1 NLTPLVGLEELMSGNHPEIRPGSFHGLSSLLKLVWVNSQVSLIERNAFDGLASLVELN 60

Qy 230 LAHNNTLLPHDLFTPLHHLRIHLHHPNWCNDILWLSWIKDMAPSNTACCARNTP 289
Db 61 LAHNNTLLPHDLFTPLHHLRIHLHHPNWCNDILWLSWIKDMAPSNTACCARNTP 120

Qy 290 PNLGKYGELDONVETCYAPVIVPEPPADLVNTEGMAELKCRASTSLTSVSWITPNGTV 349
Db 121 MHMKRYLVEVDQAQFCSAPFIMDAPKDLNISERDAELKCR-TTPMSSVWLLPNGTV 179

Qy 350 MTHGAYKVRIVASDGLTNFTNVQDTGMVTCMVSNGVNTASATLVNTAA--TTTPE 407
Db 180 LSHASRHPRIISVLDGTLNFSRVLLIDTGVVTCMVTVNAGNSASAYLVNVSSEALNTPNF 239

Qy 408 SYFSTVTVTEWPEQDEARTDNNVGPVVDVETNTVNTVTSLTQSRSTREKTTPIPTVD 467
Db 240 SFFTVTVTEWPEQDEARTDNNVGPVVDVETNTVNTVTSLTQSRSTREKTTPIPTVD 298

Qy 468 INSGI-PGIDEV 478
Db 299 TTDKMQTSLDEV 310

RESULT 6
Q9Bz20 PRELIMINARY; PRT; 606 AA.
AC Q9Bz20
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE BA438B23.1 (NEURONAL LEUCINE-RICH REPEAT PROTEIN) (CDNA FLJ31810 FIS,
DE CLONE NT2R12009289, WEAKLY SIMILAR TO CARBOXYPEPTIDASE N 83 KDA
DE CHAIN).
DE BA438B23.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Babbage A.;
RA Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA Ishibashi T., Kanehori K., Yosida M., Watanabe S., Ishida S., Ono Y.,
RA Hotuta T., Hiraoka S., Murakawa K., Takiguchi S., Kusano J.,
RA Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Takahashi-Fujii A., Oshina A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Naganari K., Masuno Y., Nagai K., Isogai T.;
```

```
*NEDO human cDNA sequencing project.*;
Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL353746; CAC22713.1; -.
DR EMBL; AK056372; BAB71167.1; -.
DR InterPro; IPR003599; Ig; 1.
DR InterPro; IPR003598; Ig_C2.
DR InterPro; IPR003600; Ig_Like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Cterm.
DR InterPro; IPR003591; LRR_Typ.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00560; LRR; 11.
DR Pfam; PF01463; LRRCT; 1.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00409; IG; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00410; IG_Like; 1.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_Typ; 10.
KW Immunoglobulin domain.
SQ SEQUENCE 606 AA; 68065 MW; CB608E281B066B9D CRC64;

Query Match 22.48; Score 565.5; DB 4; Length 606;
Best Local Similarity 29.38; Pred. No. 2.6e-34;
Matches 157; Conservative 79; Mismatches 196; Indels 103; Gaps 11;

Qy 3 CPSCVCSNQSFQVCKVIRKKNLREVPDGTSTNRLNLHENOIQIKVNSFKHLRHELIQ 62
Db 28 CPARCECSAQNKSVCCHRRLLIAIEGPIETKILDKSLKSLVNPPEEFTSYPLLEID 87

Qy 63 LSRNHIETIEGANGLANLTLEFONRLTTPNGAFVYLSKLKELWLRNNPSTPSY 122
Db 88 LSDNIIANVEGAFNNFLNLSRLKGNRLKLVPLGVFTGLSNLTSLDISENKIVLLDY 147

Qy 123 AFNRIPSLRLDLGELKRLSVISEGAPEGLSNRLYLANLMLNLRIP----- 169
Db 148 MFQDLHNLKSLVEGD-NDLVYISHRAFSGLLSLQLTLEKCNLTAVPTPEALSHRLSL 206

Qy 170 -----NLTP----- 173
Db 207 HLKHLINNMPPVAFKRLFLHLKLEIDYWPLLDMPANSLYGLNLTSLSTNTNLSTVPF 266

Qy 174 -----LKLDELDSGNHLSAIRPGSFGGLMHLQKLWMIQSQIQVIERNAFNLQSLVEI 228
Db 267 LAFKHLVYLTHLNLSYNPISTIEAGMFSDLRLQELHIVGAQLRTIEPHSFQGLRFLVL 326

Qy 229 NLAHNNLTLLPHDLFTPLHHLRIHLHHPNWCNDILWL-----SMWIKDMAPSNTACC 283
Db 327 NVSONLLETLEENVFSPRALEVISINNPLACOCRLNLIQROPTLQFGQQP----- 380

Qy 284 ARCTPNPLKGRYIGELDON----YFTCYAPVIVPEPPAD-LNVTEGMAELKCRASLTSLT 338
Db 381 -MCAGPTDIRERSEKDFHSTALSIFYFTCKPKIREKKLQLHLVDEGQTVQLECSADGDPQ 439

Qy 339 SV-SWITPNG---TWMTGAYKVRIVASDGLTNFTNVQDTGMVTCMVSNGVNTAS 394
Db 440 PVISWVTPRRRFTITKNG-----RATVLGDTGLEIRFAQQDQDSGMVYCIASNAAGNDTFT 495

Qy 395 ATLNVTAATTPFSYFSTVTVTEWPEQDEARTDNNVGPVVDVETNTVNTVTSLS 449
Db 496 ASLVKGFASDRFLYANRTPMWTDSNDTISNGTNAN---TFLDLKTLILYSTAM 547

RESULT 7
Q96FE5 PRELIMINARY; PRT; 614 AA.
ID Q96FE5
AC Q96FE5
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
```

01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
UNKNOWN (PROTEIN FOR MGC:17422).
Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRIN, AND ANAPLASTIC OLIGODENDROGLIOMA WITH 1P/19Q LOSS;
RA Strausberg R.; and 2001 to the EMBL/GenBank/DBJ databases.
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC011057; AAH11057.1; -
SQ SEQUENCE 614 AA; 69145 MW; EFD967E3B716698D CRC64;

Query Match 22.4%; Score 565.5; DB 4; Length 614;
Best Local Similarity 30.3%; Pred. No. 2.7e-34;
Matches 162; Conservative 74; Mismatches 202; Indels 93; Gaps 12;

QY 3 CPVSCSNQFSKVICVRKKNLREVPDGISTNRLNLHENOIOIKVNSFKHLRHLILQ 62
DB 36 CPPCECSAQRVLCRKRFVAVPEGIPETRLDLGKRIKTLNODERFASPHLELE 95
QY 63 LSRNHRTIEIGAPNGLANLTLEFDNRLTTPNGAFVYLSKLKELWLNRPPIESIPSY 122
DB 96 LNEINISAVEPGAFNNFLNLTGLRSNRLKLIPLGVFTGLSLNLTLDISENKIVILLY 155
QY 123 AFNRIPLRLDLGELKRLSYISEGAFEGISLNRYLNAMCNLREIP- 169
DB 156 MFQDLYNKLSEVGD-NDLYIISHRAFSGLSLEQLTLEKCNLTSTPTALSHLHGLIVL 244
QY 170 -----NLTP- 173
DB 215 RLRLHNLNARIDYFKRLYRLKLVLEISHWPYLDTPNCLYGLNLTSLSTHCLNTAVPY 274
QY 174 -----LIKLELDLSGNHLSAIRPGSQGLMHLOKLMWIOSQIOVIERNAPDNLSLVEI 228
DB 275 LAVRHLYLRLNLSYNPSTIEGSMHLHLLRQEIQLVGGQAVVEFYAFRGLNLYRLV 334
QY 229 NLAHNNLTLLPHDLFTPLHLERHLHHPNWCNCDTLWL- - -SWWIKDMAPSNTACCAR 285
DB 335 NVSGNQTLLEESVFHSGVGNLETLILDSNPLACDRLWVFRWRRLNFRNQPT- 389
QY 286 CNTPNLKGRIYGE- - -LDQNYFTCYAPVIVEPPA-DLNVTEGMAAELKRA-STSLTS 339
DB 390 CATPEFVGQKEKDFPDVLLPNYFTCRRARIRDRKAQOVFDEGHTVQFVCRADGPPPA 449
QY 340 VSWITPNTGTVTHGAYKRVIAVSDGTLNFTNTVQDTGMVTCMVSNGVNTTASATLVN 399
DB 450 ILWLSPRKHLVS-AKSNGLRTVFPDGTLEVRYAQVQDNGTYLCTAANAGGNDSPAHLV 508
QY 400 TA-ATTTTFYSFVTVETMEPSQDEARTDNNVGPVVDWETNTTSL 449
DB 509 RSYSPDHPHQPNTFAFIPNQPGEGEANSTRATV-PPF-FDIKTLIIATTM 557

RESULT 8
Q9N008 PRELIMINARY; PRT; 614 AA.
AC Q9N008;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 69.2 KDA PROTEIN.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=CEREBELLUM CORTEX;

Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
Suzuki Y., Sugano S., Hashimoto K.;
RT Isolation of full-length cDNA clones from macaque brain CDNA
libraries.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB046639; BAB03557.1; -
DR HSSP; P23945; 1XUN.
DR InterPro; IPR003598; Iq_c2.
DR InterPro; IPR003006; Iq_MHC.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003592; LRR_Out.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00047; Iq; 1.
DR Pfam; PF00560; LRR; 9.
DR Pfam; PF01463; LRRCT; 1.
DR Pfam; PF01462; LRRNT; 1.
DR SMART; SM00408; Igc2; 1.
DR SMART; SM00370; LRR; 5.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_Typ; 1.
KW Hypothetical protein; Immunoglobulin domain.
SQ SEQUENCE 614 AA; 69187 MW; BA6C8BC7C993BE9A CRC64;

Query Match 22.2%; Score 560.5; DB 6; Length 614;
Best Local Similarity 30.3%; Pred. No. 6.3e-34;
Matches 161; Conservative 74; Mismatches 203; Indels 93; Gaps 12;

QY 3 CPVSCSNQFSKVICVRKKNLREVPDGISTNRLNLHENOIOIKVNSFKHLRHLILQ 62
DB 36 CPPCECSAQRVLCRKRFVAVPEGIPETRLDLGKRIKTLNODERFASPHLELE 95
QY 63 LSRNHRTIEIGAPNGLANLTLEFDNRLTTPNGAFVYLSKLKELWLNRPPIESIPSY 122
DB 96 LNEINISAVEPGAFNNFLNLTGLRSNRLKLIPLGVFTGLSLNLTLDISENKIVILLY 155
QY 123 AFNRIPLRLDLGELKRLSYISEGAFEGISLNRYLNAMCNLREIP- 169
DB 156 MFQDLYNKLSEVGD-NDLYIISHRAFSGLSLEQLTLEKCNLTSTPTALSHLHGLIVL 214
QY 170 -----NLTP- 173
DB 215 RLRLHNLNARIDYFKRLYRLKLVLEISHWPYLDTPNCLYGLNLTSLSTHCLNTAVPY 274
QY 174 -----LIKLELDLSGNHLSAIRPGSQGLMHLOKLMWIOSQIOVIERNAPDNLSLVEI 228
DB 275 LAVRHLYLRLNLSYNPSTIEGSMHLHLLRQEIQLVGGQAVVEFYAFRGLNLYRLV 334
QY 229 NLAHNNLTLLPHDLFTPLHLERHLHHPNWCNCDTLWL- - -SWWIKDMAPSNTACCAR 285
DB 335 NVSGNQTLLEESVFHSGVGNLETLILDSNPLACDRLWVFRWRRLNFRNQPT- 389
QY 286 CNTPNLKGRIYGE- - -LDQNYFTCYAPVIVEPPA-DLNVTEGMAAELKRA-STSLTS 339
DB 390 CATPEFVGQKEKDFPDVLLPNYFTCRRARIRDRKAQOVFDEGHTVQFVCRADGPPPA 449
QY 340 VSWITPNTGTVTHGAYKRVIAVSDGTLNFTNTVQDTGMVTCMVSNGVNTTASATLVN 399
DB 450 ILWLSPRKHLVS-AKSNGLRTVFPDGTLEVRYAQVQDNGTYLCTAANAGGNDSPAHLV 508
QY 400 TA-ATTTTFYSFVTVETMEPSQDEARTDNNVGPVVDWETNTTSL 449
DB 509 RSYSPDHPHQPNTFAFIPNQPGEGEANSTRATV-PPF-FDIKTLIIATTM 557

RESULT 9
Q9D1T0 PRELIMINARY; PRT; 614 AA.
ID Q9D1T0
AC Q9D1T0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)

Query Match		16.4%	Score 413.5;	DB 4;	Length 1094;
Best Local Similarity		28.3%	Pred. No. 1.5e-22;		
Matches 127;		Conservative	78;	Mismatches 170;	Indels 73; Gaps 15;
Qy	37	LNLHENOQIITIKVNSFKHLRHLLEILQLSRNHIRTETIGAFNGLANLNTLELFDNRLTIP	96		
Db	216	LDLNRNRIRLTLEGLTFOGLNSLEVLFKLQRNNISKLTDCGAFWGLSKMHVHLHLEYNLSLVEVN	275		
Qy	97	NGAFVYLSKKELRLRNPIESIPS-----YAFNRIPSLRRLDLGELKRL--	141		
Db	276	SGSLYGLTALHQLHLSNNSIARIHRKGSFCQKLHELVLSPNNLTRDEESLAELSLSLV	335		
Qy	142	-----SYISEGAFEGLSNRLYLNLAMCNLREIPNLPLIKDELDELDSGNHLSAIRPG	193		
Db	336	LRLSHNSISHIAEGAFKGLSLRVLD-----LDHNEISGTIEDT--SG	376		
Qy	194	SFOGLMHLQKLWMIQSIOQVIERNAFQNLQSLVEINLAHNNLTLLPHDLFTPLHLHERIH	253		
Db	377	AFSGLDLSLKLTLFGNKIKSVAKRAFSGLEHNLGGNAIRSVQFDADFVKMKNLRELH	436		
Qy	254	LHHNPWCNCDILNLSMWI--KDMAPSNWACCARCNTPPNLKGRYIGELDONYETC---Y	308		
Db	437	ISSDSFLDCOLKWLPPWLIIGRMLOAFVTATCAH---PESLKGSIFSVPPESEFVCDDEL	493		
Qy	309	APVIVEPPADLNVTEGMAAELKCRASLTS---VSWITPNGTVMTGAYKVIAY-LSD	364		
Db	494	KPQIITQPETTMAMVGKDIRETCSAASSSSPMTFAKKDN-EVLTNADMENFVHVHAQD	552		
Qy	365	G-----TLNFTNVTVDGTGMYTCMVNSVGNL---TASATLNVTAA-TTTPFSYFSTV	413		
Db	553	GEVMEYTTILHLRQVTEGHEGRYQCVTITNHFGSTYSHKARLTNNVLPSTKTPHD----	608		
Qy	414	TVETMEPSODEARTDNNVGPVVDWE	441		
Db	609	TIRTTMARLECAATGH---PNQIAWO	633		

Search completed: August 26, 2002, 15:50:24
Job time: 246 sec

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	421	16.7	708	4	US-09-131-648-2	Sequence 2, Appli
2	402.5	15.9	1091	3	US-08-986-485-5	Sequence 5, Appli
3	393	15.6	1101	3	US-08-986-485-2	Sequence 2, Appli
4	386	15.3	1525	3	US-09-191-647-2	Sequence 2, Appli
5	386	15.3	1525	4	US-09-540-245A-2	Sequence 2, Appli
6	386	15.3	1525	4	US-09-540-153-2	Sequence 2, Appli
7	372	14.7	1523	4	US-09-182-024A-2	Sequence 2, Appli
8	362	14.3	673	4	US-09-063-950-2	Sequence 2, Appli
9	354	14.0	649	4	US-09-188-930-305	Sequence 305, App
10	350	13.9	1480	3	US-09-191-647-7	Sequence 7, Appli
11	350	13.9	1480	4	US-09-540-245A-7	Sequence 7, Appli
12	350	13.9	1480	4	US-09-540-153-7	Sequence 7, Appli
13	350	13.9	1480	5	PCM-US91-09055-2	Sequence 2, Appli
14	347	13.7	1480	4	US-09-182-024A-5	Sequence 5, Appli
15	329.5	13.1	560	3	US-08-592-500-2	Sequence 2, Appli
16	329.5	13.1	560	3	US-08-195-006-2	Sequence 2, Appli
17	329.5	13.1	560	4	US-09-063-950-4	Sequence 4, Appli
18	329.5	13.1	560	5	PCM-US94-07644A-2	Sequence 2, Appli
19	329.5	13.1	605	4	US-09-063-950-5	Sequence 5, Appli
20	314.5	12.5	605	1	US-08-190-802A-49	Sequence 49, Appl
21	314.5	12.5	605	4	US-08-477-346-49	Sequence 49, Appl
22	314.5	12.5	605	4	US-08-473-089-49	Sequence 49, Appl
23	309	12.2	603	1	US-08-190-802A-50	Sequence 50, Appl
24	309	12.2	603	4	US-08-477-346-50	Sequence 50, Appl
25	309	12.2	603	4	US-08-473-089-50	Sequence 50, Appl
26	303.5	12.0	353	3	US-08-986-485-6	Sequence 6, Appli
27	302	12.0	282	1	US-08-442-063A-45	Sequence 45, Appl

Db 328 NAFRLPKLESLMNSALYHGTIESLPNLKEISHSNPIRCDCVIRWMNNKINIR 387
QY 274 DMASNTACCARCNTPNLGRYIGELD-QNYFTCYAPVIVP--PADLNVTEGMAAELK 330
Db 388 FMEPDSLCV---DPPEFQONVRQVHFDRDMMIEICLPIAPEFSPSNLNVAGSVSFH 443
QY 331 CRASTS-LTYSVITPNTGTYMTHGAYKVRIVAVLSGTLNFTNTVQDTGMYTCMVNSVYG 389
Db 444 CRATAEPOPEIYITPSGQKLLPNTLTKDFYVHSEGLTNDVTPRKGGGLYTCIATNLVG 503
QY 390 NTTASATLV 399
Db 504 ADLKSVMIKV 513

RESULT 2

US-08-986-485-5
; Sequence 5, Application US/08986485
; Patent No. 6046030
; GENERAL INFORMATION:
; APPLICANT: WU, SHUJIAN
; APPLICANT: SWEET, RAYMOND
; APPLICANT: TRUNEH, ALEMSEGED
; TITLE OF INVENTION: A HUMAN LIG-1 HOMOLOG (HLIG-1)
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/986,485
; FILING DATE: 08-DEC-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/059,448
; FILING DATE: 22-SEP-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70264
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846189
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1091 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-986-485-5

Query Match 15.9%; Score 402.5; DB 3; Length 1091;
Best Local Similarity 24.1%; Pred. No. 5.1e-25;
Matches 134; Conservative 77; Mismatches 168; Indels 177; Gaps 14;

QY 37 LNLHENOIQIKVNSFKLRLHLEIQLSRNHIRTIEIGAFNGLANLTLFNLFTTIP 96
Db 218 LDLNRRNRIEGLTFQGLDLSLEVLQRNISRITDGAFWGLSKMHVHLHLEYNLSVEVN 277
QY 97 NGAFVYLSKLKELMLRNPNSTPS-----YAFNRIPIRRLDLGELKRL-- 141

Db 278 SGLSYGLTALHQLHLSNNSISRIQDQMSFCQKHELILSFNNLTLRDEESLAELSLSI 337
QY 142 -----SYISEGAFGLSNLRYLNLAMCNLRIPNLTPLIKDELDSLGNHLSATRP 193
Db 338 LRLSHNAISHIAEGAFGLKSLRVL-----LDHNEISGTIEDT--SG 378
QY 194 SFQGLMHLOKLMITQSOIQVIERNAFDNLOSLEINLAHNNLTLPHDLFTPLHLERIH 253
Db 379 AFTGLDNLKLTLPNGKIKSVAKRAFGSGLSEHNLGENAIRSVQDFAFKMKNLKELY 438
QY 254 LHHNPNWCNDIWLWSWI--KDMAPNTACCARCNTPNLKGRI--GELDONY----- 304
Db 439 ISSSEFLDCOLKWLPPWLMGRMLQAFVTATCAH---PESLKGQISFVLPSDFVCDP 495
QY 305 -----FTCYA----- 309
Db 496 KPQIITQPETTMAYVGKDIRFTCSAASSSSSPMTFAWKONEVLANADMENFAHRAQDG 555
QY 310 -----PVIVEPPADLNVTEG 324
Db 556 EVMEYTTILHLRHVTFGHEGRYQCIITNHRGSTYSHKARLTVNVLPSETKIPHDIAITG 615
QY 325 MAELKCRASLT--SVSWITPNTGTYMTHGAYKVRIVAVLSGTLNF--TNVTQDTGMYTC 382
Db 616 TTRLECAATGHPNPQIAWQKDGTDFF--AARERRMHVMPDDDDVFFITDKIDDGMVYSC 674
QY 383 MVSNSVGNNTASATLVNTAATTTTFSYFVTVTETMPSODEARTDNNV-----GP 434
Db 675 TAQNSAGSVSANATLVLE-----TPSLAVPLEDRVVTVGETVAFOCKATGSP 722
QY 435 TPVVDWETNTVNTSLT 450
Db 723 TPRTWLKGGRLSLT 738

RESULT 3

US-08-986-485-2
; Sequence 2, Application US/08986485
; Patent No. 6046030
; GENERAL INFORMATION:
; APPLICANT: WU, SHUJIAN
; APPLICANT: SWEET, RAYMOND
; APPLICANT: TRUNEH, ALEMSEGED
; TITLE OF INVENTION: A HUMAN LIG-1 HOMOLOG (HLIG-1)
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/986,485
; FILING DATE: 08-DEC-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/059,448
; FILING DATE: 22-SEP-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70264
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846189
; INFORMATION FOR SEQ ID NO: 2:


```
Qy 161 AMCN-----LRETP-NL-----TPL 174
Db 271 PSCNANSISCPSPCTCSNNIVDCRGKGLMEIPANLPEGIVEIRLEQNSIKAIPAGAFQY 330
Qy 175 IKLDELDSGNHLSAIRPGSFQ-----GLMHQKLMHTSQ 210
Db 331 KKLKRIDISKNOISDIAPAFQGSQKSLTSVLYGNKITEIAKGLFDGLVSLQLLLLNANK 390
Qy 211 IQVTERNAFDLQSLVEINLAHNNLTLLPHDLFTPLHHLRIHLHHPNWCNCDILWLSW 270
Db 391 INCLRVNTFQDLQNLNLLSLVDNKLQYISKGLFAPLQSIQTLHLAQNPFVCDCHLKLWAD 450
Qy 271 WIDMAPSNTACCARCNTPPNLKGRYIGELDONFTCYAPVIVEPPADNLVTEGMAELK 330
Db 451 YLQD-NPIETS-GARCSSPRLANKRISQIKSKFRCSGSEDYRSFESCFMDLVCPCK 508
Qy 331 CRASLTSLTSWITPNCVTWTHGAYK-VRI-----AVLSDGTLANFTNTVV-QDTGMY 380
Db 509 CRC-----EGTIVDCSNOKLVRIPSHLPEYVTDLRNDNEVSLVLEATGIF 553

RESULT 8
US-09-063-950-2
; Sequence 2, Application US/09063950C
; Patent No. 6225085
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NOVEL LRSG PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: MEI-019
; CURRENT APPLICATION NUMBER: US/09/063.950C
; CURRENT FILING DATE: 1998-04-21
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 673
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-063-950-2

Query Match 14.38; Score 362; DB 4; Length 673;
Best Local Similarity 29.98; Pred. No. 6e-22;
Matches 115; Conservative 56; Mismatches 150; Indels 64; Gaps 11;

Qy 1 QTCPSVCSNQSFKVICVRKNLREVPDGISTNTRLLNLHENQIOIKVNSFKHLRLHLEI 60
Db 22 QGCPSGCQCS-QPQTVCTARQGTTPRDPVDPDTVGLVVFENGITMLDAGSFAGLPGLQ 80
Qy 61 LQLSRNHRTIEIGAFNGLANLTLEFDNRLTTPINGAFVYLSKLKELWLNRNPIESIP 120
Db 81 LDLSQNIASLPQSGVQPLANLSNLDLTANRLHEITNETFRGLRLRLYLKGNRIHQ 140
Qy 121 SYAFN-----RTPSLRRDLGLKELRLSYISEGAFGLSLNRLYL 159
Db 141 PGAFDTLRLLEKLQDNELRALPLRLPLRLDLSNLSLAL--EPGILDANVEALR 198
Qy 160 LAMCNLRIPN--LTPLIKDELDSGNHLSAIRPGSFOGLMHQKLMWI--OSQIOVTER 216
Db 199 LAGLQQLDDEGLSRLNHLHDVSDNQLERV-PPVIRGLRGLTRLRAGNTRIAQLRP 257
Qy 217 NAFDNLQSLVEINLAHNNLTLLPHDLFTPLHHLRIHLHHPNWCNCDILWLSWKID-- 274
Db 258 EDLAGLAALQELDVSNLSLQALPGDLSGLFPLRLRLAAARNPENCVCPLSFWGPMVRESH 317
Qy 275 ---MAPSNTACCARCNTPPNLKGRYIGELDONFTCYA-----PVIVEPPADLN 320
Db 318 VTLASPEET-----RCHFPKKNAGRLULLLDYADFGCPATTTTATVPTTRPVVREPTA--- 370
Qy 321 VTEGMAELKCRASLTSLTSWITP 345
Db 371 -----LSSSL-APTWLSP 382
```

```
RESULT 9
US-09-188-930-305
; Sequence 305, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188.930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 305
; LENGTH: 649
; TYPE: PRT
; ORGANISM: Mouse
US-09-188-930-305
```

```
Query Match 14.08; Score 354; DB 4; Length 649;
Best Local Similarity 24.68; Pred. No. 2.7e-21;
Matches 136; Conservative 88; Mismatches 202; Indels 126; Gaps 20;

Qy 1 QTCPSVCSNQSFKVICVRKNLREVPDGISTNTRLLNLHENQIOIKVNS-FKHL----- 55
Db 29 KSCPVCRCDAGF--IYCNDRSLTSIPGIPEDATTLVQNQINNNGIPSDLKLLKVVQ 86
Qy 56 -----RHLEILOLSNRHIRTIEIGAENGLANLTLELNFNRLT--TIP 96
Db 87 RIYLYHNSLDEFPPTNLPYIVRELHLQENNIITTYVDSKIPYELDELDDONSVAVSIE 146
Qy 97 NGAFVYLSKLKELWLNRNPIESIPSAFNRIPLSRRLDLGELKRLSYISEGAFGLSLNLR 156
Db 147 EGAFDSNYLRLLFLSRNHLSTIPGGLPRTIEELRLDD---NRISTISSPSLHGLTSLK 202
Qy 157 YLNAMCNLRIPNLT-----LIKDELDSGNHLSAIRPGSFOGLMHQKLMWIQSOI 211
Db 203 RLVLDD-GNLLNNHGLGDKVFFNLVNLTELSLVRNSLTA-APVNLPG-TSLRKLVLQDNHI 259
Qy 212 QVIERNAFDNLQSLVEINLAHNNLTLLPHDLFTPLHHLRIHLHHPNWCNCDILWLSW 271
Db 260 NRVPNFAFSYLRQLYRLDMNNSNLSNLPGIFDDLDNITQLILRNPNWYCCCKMKWRDW 319
Qy 272 IKDMAPSNTACCARCNTPPNLKGRYIGELDONFTCY----- 308
Db 330 LQSLPVKYNVRGLMQCAPEKVRGMAIKDLSAELFDCGSGIVSTIQTITAITPNTAYPAQ 379
Qy 309 ---APVIVEPP-----ADLNVTEGMAELKCRASLTSLTSWITPNCVTWTHGAYKVR 358
Db 380 QMPAPVTQPDIKNPKLIKQDQRTTGSPRK-----TLITVKSVPD---TIHISWRIA 430
Qy 359 IAVLSDGTLANFTNTVQDTGMVTCMVNSVGNVTASATLNVTAATTPFSYFST----- 412
Db 431 LPM-----TALRLSWL--KLHSPAFGSIITIVTGERSEYLVTALEPES 473
Qy 413 ---VTVETMEPSQ-----DEARTDNNVGTPV-VDWETT-----NVVTSLTPOSTRSTEK 459
Db 474 PYRCVMPMETSNLYLFDE-----TPVCIETQAPLRMYNPTITLNRQEKEPKYK 523
Qy 460 TFTPVTDTNSG 471
Db 524 NPNLPALAAIIG 535
```

```
RESULT 10
US-09-191-647-7
; Sequence 7, Application US/09191647
```


US-09-540-153-7

```

Query Match      13.9%; Score 350; DB 4; Length 1480;
Best Local Similarity 23.9%; Pred. No. 1.9e-20;
Matches 109; Conservative 65; Mismatches 152; Indels 130; Gaps 9;

QY 2 TCPSVCSNOFSKVICVRKNLREVPDGIISTNTRLLNLHENQIOIKVNSFKHLRHLEIL 61
Db 294 SCPPHRCADGI--VDCREKSLTSPVTPDPDPTDVTLEQNFITELPKFSFRRRLRI 351
QY 62 QLSNRHRTTEIGAFNGLANLTELFDNRLLTTPNGAFVYLSKLKELWLRNNPIESIPS 121
Db 352 DLSNNISRTAHDAISGLKQLTTLVLYGNKIKDLPKGVFKGLSLRLLNANEISCIRK 411
QY 122 YAFNRIPSLRLDLGELKRLSYISEGAFGLSNLRYLNLM-----CNLR----- 166
Db 412 DAFRDLHSLLSLYD--NNIQSLANGTFDAMKSMKTVHLAKNPFICDNLRLWADYLHN 470
QY 167 -----EIPNLTP-----IKLD----- 166
Db 471 PIETSGARCESPKRMHRRRIETSLREKFKCSWGELRMKLSGECRMDSDCPAMCHCEGTV 530
QY 167 -----EIPNLTP-----IKLD----- 166
Db 531 DCTGRRLEKEPRDIPPLHTTELLNDNLGRISDGLFGRPLHLVKLEKRNQLTGIEPNA 590
QY 195 FQGLMHLOKLMMIQSOQIOVIERNAFDNLQSLVEINLAHNNLTLLPHDLFTPLHLRIHL 254
Db 591 FEGASHIOELQGENKIKEISNKMFLGLHOLKTLNLDNQISCVMPGSFEHLNLSLNL 650
QY 255 HHNPWNCNCDILWLSWIKWDMAPSNTACCRCNTPPNKLGRYIGELDNQFTCYAPVIVE 314
Db 651 ASNPFCNCHLAWFAECVRKKSLLNGA--ARCGAPSKVRDQIKDLPHSEFKCSS----- 703
QY 315 PPADLNVTEGMAELKCRASLTLSVSWITPNTGTVM 350
Db 704 -----ENSEGCLGDGYCPPSCTCT-----GTVV 726

RESULT 13
PCT-US91-09055-2
; Sequence 2, Application PCT/US9109055
; GENERAL INFORMATION:
; APPLICANT: Rothberg, Jonathan Marc and Artavanis-Tsakonas, Spyridon
; TITLE OF INVENTION: Purified SLIT protein and Sequence Elements Thereof
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Yale University
; ADDRESS: Office of Cooperative Research
; STREET: 246 Church Street
; STREET: Suite 401
; CITY: New Haven
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06510
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 800 Kb storage
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 6.0.5
; SOFTWARE: Microsoft Word 4.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/09055
; FILING DATE: 19911127
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/624,135
; FILING DATE: 7-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Barth, Richard J.
; REGISTRATION NUMBER: 28,180
; REFERENCE/DOCKET NUMBER: 900964/RSB
; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: (212) 972-1400
; TELEFAX: (212) 370-1622
; TELEX: 236268
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1480 amino acids
; TYPE: AMINO ACIDS
; TOPOLOGY: Linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: signal sequence
; LOCATION: 1 to 36
; IDENTIFICATION METHOD: similarity to other signal
; OTHER INFORMATION: Directs Export
; NAME/KEY: Four Flank-LRR-Flank domains
; LOCATION: 37 to 910
; IDENTIFICATION METHOD: Array of Flank-LRR-Flank
; OTHER INFORMATION: mediates adhesive events
; NAME/KEY: Tandem EGF-like repeats
; LOCATION: 911 to 1150
; IDENTIFICATION METHOD: similarity to tandem EGF-like
; OTHER INFORMATION: protein-protein interactions
; NAME/KEY: 7th EGF-like repeat
; LOCATION: 1353 to 1393
; IDENTIFICATION METHOD: similarity to epidermal growth
; OTHER INFORMATION: Involvement in receptor-ligand
; NAME/KEY: Alternative splice segment
; LOCATION: 1394 to 1404
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: developmentally regulated
; NAME/KEY: COOH-terminal region
; LOCATION: 1405 to 1480
; IDENTIFICATION METHOD: experimental
; PCT-US91-09055-2

Query Match      13.9%; Score 350; DB 5; Length 1480;
Best Local Similarity 23.9%; Pred. No. 1.9e-20;
Matches 109; Conservative 65; Mismatches 152; Indels 130; Gaps 9;

QY 2 TCPSVCSNOFSKVICVRKNLREVPDGIISTNTRLLNLHENQIOIKVNSFKHLRHLEIL 61
Db 294 SCPPHRCADGI--VDCREKSLTSPVTPDPDPTDVTLEQNFITELPKFSFRRRLRI 351
QY 62 QLSNRHRTTEIGAFNGLANLTELFDNRLLTTPNGAFVYLSKLKELWLRNNPIESIPS 121
Db 352 DLSNNISRTAHDAISGLKQLTTLVLYGNKIKDLPKGVFKGLSLRLLNANEISCIRK 411
QY 122 YAFNRIPSLRLDLGELKRLSYISEGAFGLSNLRYLNLM-----CNLR----- 166
Db 412 DAFRDLHSLLSLYD--NNIQSLANGTFDAMKSMKTVHLAKNPFICDNLRLWADYLHN 470
QY 167 -----EIPNLTP-----IKLD----- 166
Db 471 PIETSGARCESPKRMHRRRIETSLREKFKCSWGELRMKLSGECRMDSDCPAMCHCEGTV 530
QY 167 -----EIPNLTP-----IKLD----- 166
Db 531 DCTGRRLEKEPRDIPPLHTTELLNDNLGRISDGLFGRPLHLVKLEKRNQLTGIEPNA 590
QY 195 FQGLMHLOKLMMIQSOQIOVIERNAFDNLQSLVEINLAHNNLTLLPHDLFTPLHLRIHL 254
Db 591 FEGASHIOELQGENKIKEISNKMFLGLHOLKTLNLDNQISCVMPGSFEHLNLSLNL 650
QY 255 HHNPWNCNCDILWLSWIKWDMAPSNTACCRCNTPPNKLGRYIGELDNQFTCYAPVIVE 314
Db 651 ASNPFCNCHLAWFAECVRKKSLLNGA--ARCGAPSKVRDQIKDLPHSEFKCSS----- 703
QY 315 PPADLNVTEGMAELKCRASLTLSVSWITPNTGTVM 350
Db 704 -----ENSEGCLGDGYCPPSCTCT-----GTVV 726

```

RESULT 14
US-09-182-024A-5
; Sequence 5, Application US/09182024A
; Patent No. 6342370
; GENERAL INFORMATION:
; APPLICANT: Connolly, Timothy
; APPLICANT: Rajput, Bhanu
; TITLE OF INVENTION: Human Slit Polypeptide and Polynucleotides Encoding
; TITLE OF INVENTION: Same
; FILE REFERENCE: 640100-271
; CURRENT APPLICATION NUMBER: US/09/182,024A
; CURRENT FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: 60/063,946
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/096,420
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1480
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-182-024A-5

Query Match 13.7%; Score 347; DB 4; Length 1480;
Best Local Similarity 23.9%; Pred. No. 3.3e-20;
Matches 109; Conservative 65; Mismatches 152; Indels 130; Gaps 9;
QY 2 TCPSVCSNOFSKVICVRKNLRVDPGISTNTLLNLHENOIQIIVKNSFKHLRHLLEIL 61
Db 294 SCPPCRCADGI--VDCREKSLTSVPVTLDDTDDVLEQNFITELPKFSFSSFRRLRI 351
QY 62 QLSNRHIRTIEIGAFNGLANLTLEFDNRLLTTPNGAFVYLSKLKELWLRNPIESIPS 121
Db 352 DLSNNISRIAHDALSGKQLTLLVLYGNKIKDLPFGVFKGLSLRLLLNANEISCIK 411
QY 122 YAFNRIPSLRRDLGELKRLSYISEGAFEGLSNRLRYNLAM-----CNLR----- 166
Db 412 DAFRDLHSLSLSYD--NNIQLANGTFDAMKSMKTVHLAKNPFCIDCNLRWLADYLHN 470
QY 167 ----- 166
Db 471 PIETSGARCESPKRMHRRRIEESREKFKGWSGELRMKLSGECRMDSDCPAMCHCGTV 530
QY 167 -----ETPNTPL-----IKLD-----ELDLSGNHLSAIRPGS 194
Db 531 DCTGRRLKEIPRDIPLHTTELLNDNELGRISDGLFGRPLHLVKELKRNQLTGIEPNA 590
QY 195 FQGLMHLOKLMWISQIQVIERNAFDNLQSLVEINLAHNNLTLLPHDLFTPLHHLRIHL 254
Db 591 FEGASHIQELQGENKIKEISNKMFLGLHQLKTLNLYDNOISCVMPGSEFHLNLSLNL 650
QY 255 HHNPWNCNDILWSWIKMDASNTACCARNTPPNLKGRIYIGELDQNTFTCYAPVIVE 314
Db 651 ASNPFNCNHLAWFAECVRKSLNGGA--ARCGAPSKVRVDQIKDLPHSEFKSS----- 703
QY 315 PPADLVNTEGMAAEKLCRASTLSITSVSWITPNGTVM 350
Db 704 -----ENSEGLGDGYPCTCT-----GTIV 726

RESULT 15
US-08-592-500-2
; Sequence 2, Application US/08592500
; Patent No. 6005089
; GENERAL INFORMATION:
; APPLICANT: Lanza, Francois
; APPLICANT: Phillips, David R.
; APPLICANT: Cazenave, Jean-Pierre
; TITLE OF INVENTION: Platelet Glycoprotein V Gene and Uses
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: 379 Lytton Avenue
CITY: Palo Alto
STATE: California
COUNTRY: US
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,500
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/089,455
FILING DATE: 09-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 12418-28
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 560 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-592-500-2

Query Match 13.1%; Score 329.5; DB 3; Length 560;
Best Local Similarity 29.7%; Pred. No. 2.4e-19;
Matches 83; Conservative 43; Mismatches 104; Indels 49; Gaps 3;
QY 22 NLREVDPG-----ISTNRLN-----LHENOIQIIVKNSFKH 54
Db 181 NLHLPKGLLGAQAKLERLLHNSNRLVSLDGLNSLGCALTTELQFHRNHSIAPGAFDR 240
QY 55 LRHLLEILQLSNRHIRTIEIGAFNGLANLTLEFDNRLLTTPNGAFVYLSKLKELWLRN 114
Db 241 LPNLSLTLNRHLAFPLSALFLHSHNLTLTLFENPLAELPGVLFGEMLGLQELWLRN 300
QY 115 PIESIPYANRIPSLRRDLGELKRLSYISEGAFEGLSNRLRYNLAMCNLRIPNLTP 174
Db 301 QLRTLPAAAFNRSLRRLYLVTLSPRLSALPQGAFOGLGELQV----- 343
QY 175 IKLDELDSGNHLSAIRPGSFQGLMHLOKLMWISQIQVIERNAFDNLQSLVEINLAHNN 234
Db 344 -----LALHNSGLTALPDGLLRGLGKLRQVSLRNRRLRALPRALFRNLSSLESYOLDH 398
QY 235 LTLPHDLFTPLHHLRIHLHNPWNCNDILWLSWIK 273
Db 399 LETLPQGVFGALPRLTEVLLGHNSWRCDGCLGPGELGLWR 437

Search completed: August 26, 2002, 15:46:38
Job time: 170 sec

